THE ANNALS
of
MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

VOLUME XVI

1945

THE ANNALS OF MATHEMATICAL STATISTICS

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore 2, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the ANNALS OF MATHEMATICAL STATISTICS, Mt. Royal & Guilford Aves., Baltimore 2, Md., or to the Secretary of the Institute of Mathematical Statistics, P. S. Dwyer, 116 Rackham Hall, University of Michigan, Ann Arbor, Mich.

Changes in mailing address which are to become effective for a given issue should be reported to the Secretary on or before the 15th of the month preceding the month of that issue. The months of issue are March, June, September and December. Because of war-time difficulties of publication, issues may often be from two to four weeks late in appearing. *Subscribers are therefore requested to wait at least 30 days after month of issue before making inquiries concerning non-delivery.*

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $5.00 per year. Single copies $1.50. Back numbers are available at $5.00 per volume, or $1.50 per single issue.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

THE ANNALS
of
MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE
OF MATHEMATICAL STATISTICS

Contents

Vol. XVI, No. 1 — March, 1945

THE ANNALS OF MATHEMATICAL STATISTICS

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore 2, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the ANNALS OF MATHEMATICAL STATISTICS, Mt. Royal & Guilford Aves., Baltimore 2, Md., or to the Secretary of the Institute of Mathematical Statistics, P. S. Dwyer, 116 Rackham Hall, University of Michigan, Ann Arbor, Mich.

Changes in mailing address which are to become effective for a given issue should be reported to the Secretary on or before the 15th of the month preceding the month of that issue. The months of issue are March, June, September and December. Because of war-time difficulties of publication, issues may often be from two to four weeks late in appearing. *Subscribers are therefore requested to wait at least 30 days after month of issue before making inquiries concerning non-delivery.*

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $5.00 per year. Single copies $1.50. Back numbers are available at $5.00 per volume, or $1.50 per single issue.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879

THE APPROXIMATE DISTRIBUTIONS OF THE MEAN AND VARIANCE OF A SAMPLE OF INDEPENDENT VARIABLES

By P. L. Hsu

The National University of Peking

1. Introduction. In this paper we shall study the mean and variance of a large number, n (a sample of size n) of mutually independent random variables:

$$\xi_1, \xi_2, \cdots, \xi_n, \tag{1}$$

having the same probability distribution represented by a (cumulative) distribution function $P(x)$. The rth moment, absolute moment, and semi-invariant of $P(x)$ are denoted by α_r, β_r, and γ_r respectively. It is assumed that for a certain integer $k \geq 3$, $\beta_k < \infty$ and that $\alpha_2 > 0$. Hence there is no loss of generality in assuming that

$$\alpha_1 = 0, \qquad \alpha_2 = 1. \tag{2}$$

The characteristic function corresponding to $P(x)$ is denoted by $p(t)$.

We put

$$\bar{\xi} = \frac{1}{n} \sum_{r=1}^{n} \xi_r, \qquad \eta = \frac{1}{n} \sum_{r=1}^{n} (\xi_r - \bar{\xi})^2 \tag{3}$$

$$F(x) = Pr\{\sqrt{n}\,\bar{\xi} \leq x\}, \qquad G(x) = Pr\left\{\frac{\sqrt{n}(\eta - 1)}{\sqrt{\alpha_4 - 1}} \leq x\right\}. \tag{4}$$

The definition of $G(x)$ implies that $\alpha_4 < \infty$ and $\alpha_4 - 1 > 0$. The case $\alpha_4 - 1 = 0$ provides an easy degenerated case which will be treated separately (section 4).

Cramér's theorem of asymptotic expansion[1] reads as follows:

Theorem 1. *If $P(x)$ is non-singular and if $\beta_k < \infty$ for some integer $k \geq 3$, then*

$$F(x) = \Phi(x) + \Psi(x) + R(x) \tag{5}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\, dy. \tag{6}$$

$\Psi(x)$ is a certain linear combination of successive derivatives $\Phi^{(3)}(x), \cdots, \Phi^{(3(k-3))}(x)$ with each coefficient of the form $n^{-\frac{1}{2}\nu}$ times a quantity depending only on $k, \alpha_3, \cdots, \alpha_{k-1}$ $(1 \leq \nu \leq k - 3)$ and

$$| R(x) | \leq Q/n^{\frac{1}{2}(k-2)} \tag{7}$$

where Q is a constant depending only on k and $P(x)$.

---

[1] H. Cramér: *Random Variables and Probability Distributions* (1937), Ch. 7. This book will be referred to as (C).

In particular, putting $k = 3$ we get that $|F(x) - \Phi(x)| \leq Qn^{-\frac{1}{2}}$ provided $P(x)$ is non-singular and $\beta_3 < \infty$. If the condition of non-singularity of $P(x)$ be removed, then Liapounoff's theorem[2] furnishes the weaker result: $|F(x) - \Phi(x)| \leq A\beta_3 n^{-\frac{1}{2}} \log n$ where A is a numerical constant.

Very recently Berry[3] succeeded in removing the factor $\log n$ from Liapounoff's theorem under no other condition than that $\beta_3 < \infty$. We state here Berry's theorem:

THEOREM 2. *If $\beta_3 < \infty$, then*

$$|F(x) - \Phi(x)| \leq \frac{A\beta_3}{\sqrt{n}} \tag{8}$$

where A is a numerical constant.

An essential step in the proof of these results is the selection of a weighting function $w(x)$ and the appraisal of the integral

$$\int_{-\infty}^{\infty} w(u)\{F(u + x) - \Phi(u + x) - \Psi(u + x)\}\, du \tag{9}$$

($\Psi \equiv 0$ when $k = 3$). In his book[1] Cramér proves Theorem 1 by taking $w(u) = \frac{1}{\Gamma(\omega)}(-u)^{\omega-1}$ when $u < 0$ and $w(u) = 0$ when

$$u \geq 0 \quad (0 < \omega < 1) \tag{10}$$

and proves Liapounoff's theorem by taking

$$w(u) = \frac{1}{\sqrt{2\pi}\epsilon} e^{-u^2/2\epsilon^2}. \tag{11}$$

On the other hand, Berry uses the following weighting function in his proof of Theorem 2:

$$w(u) = \frac{1 - \cos Tu}{u^2}. \tag{12}$$

The unfortunate selection of the function (11) accounts for the presence of the factor $\log n$ in Liapounoff's theorem.

Now Cramér's proof of Theorem 1, based on the integral (9) with $w(u)$ defined in (10), makes use of a result on that integral due to M. Riesz. A more elementary proof than this can be devised. In fact, one has only to use, with Berry, the function (12) and to adopt his elementary appraisal[4] of the integral

---

[2] (C), Ch. 7.

[3] A. C. BERRY: "The accuracy of the Gaussian approximation to the sum of independent variates." *Trans. Amer. Math. Soc.*, Vol. 49 (1941), pp. 122–136. This paper will be referred to as (B).

[4] Berry proves the inequality (in our notation):

$$\left|\int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2}\{F(x + a) - \Phi(x + a)\}\, dx\right| \leq \int_0^T \frac{(T - t)\,|f(t) - e^{-\frac{1}{2}t^2}|}{t}\, dt$$

(9) in order to obtain the proof of Theorem 1. One of our purposes is therefore to give an elementary proof of Theorem 1, without reference to the above-mentioned result due to M. Riesz. Section 2 is devoted to this work.

We ought to add that Cramér's theorem and Berry's theorem correspond to Theorems 1 and 2 for the case in which the random variables (1) do not follow the same distribution. The proof given in Section 2 is adaptable to these more general theorems when subjected to appropriate modifications; the assumption of a common distribution function for (1) is only made for the sake of convenience.

So much for the known results for the approximate distribution of $\bar{\xi}$. By a purely formal operational method Cornish and Fisher[5] obtain terms of successive approximation to the distribution function of any random variable X with the help of its semi-invariants. It is hardly necessary to emphasize the importance of turning Cornish and Fisher's formal result (asymptotic expansion without appraisal of the remainder) into a mathematical theorem of asymptotic expansion which gives the order of magnitude of the remainder. In this paper we achieve this for the simplest function of (1) next to $\bar{\xi}$, viz. the η in (3). We do not seek to remove the assumption of a common distribution for (1), as there will be no practical significance (e.g. in statistics) of η if the variables (1) do not have the same probability distribution. Section 3 is devoted to the proof of the following theorems:

THEOREM 3. *If $\alpha_6 < \infty$ and $\alpha_4 - 1 - \alpha_3^2 \neq 0$ (it cannot be negative), then*

$$|G(x) - \Phi(x)| \leq \frac{A}{\sqrt{n}} \left(\frac{\alpha_6}{\alpha_4 - 1 - \alpha_3^2} \right)^{3/2} \tag{13}$$

where A is a numerical constant.

THEOREM 4. *Let $P(x)$ be non-singular and let $\alpha_{2k} < \infty$ for some integer $k > 3$. Then*

$$G(x) = \Phi(x) + \chi(x) + R_1(x), \tag{14}$$

where $\Phi(x)$ is the function (6), $\chi(x)$ is a linear combination of the derivatives $\Phi'(x)$, $\cdots, \Phi^{(3(k-3))}(x)$ with each coefficient of the form $n^{-\frac{1}{2}\nu}$ times a quantity depending only on k and $\alpha_3, \alpha_4, \cdots, \alpha_{2k-2}$, and

---

(B), p. 128. The "appraisal" mentioned here refers to (50) which is contained in B, p. 128. But Berry's appraisal of the integral in the right-hand side of the above inequality is in default. He writes

$$\frac{\epsilon}{6} \int_0^{c/\epsilon} \left(\frac{1.1}{\epsilon} - t \right) t^2 e^{-\frac{1}{2}t^2} \, dt = \frac{1.1}{6} \sqrt{\frac{\pi}{2}} - \frac{\epsilon}{3} - \frac{1}{6} \int_{c/\epsilon}^{\infty} \left\{ (1.1 - c)t^2 + c - \frac{2c}{t^2} \right\} e^{-t^2/2} \, dt$$

(B, p. 132, line 3) whilst the last integral ought to be

$$\int_{c/\epsilon}^{\infty} \{(1.1 - c)t^2 + c - 2\epsilon t\} e^{-t^2/2} \, dt.$$

[5] E. A. Cornish and R. A. Fisher: "Moments and cumulants in the specification of distributions." (Revue de l'Institut International de Statistique (1937), pp. 1–14.)

$$|R_1(x)| \leq \frac{Q_k}{n^{\frac{1}{2}(k-2)}} \quad \text{if } k = 4, 5 \text{ or } 6 \tag{15}$$

$$|R_1(x)| \leq \frac{Q_k'}{n^{k(k-1)/(2k+3)}} \quad \text{if } k \geq 7 \tag{16}$$

where Q_k and Q_k' are constants depending only on k and $P(x)$.

It may be noticed that Theorem 3 is a "Berryian" theorem about $G(x)$, its characteristic feature being the absence of any condition on the distribution function except the two on its moments, and that Theorem 4 is a "Cramerian" theorem about $G(x)$, the characteristic feature being the assumption of non-singularity of $P(x)$ besides that $\alpha_{2k} < \infty$.

In proving these theorems we have devised a method which is applicable to getting similar results about functions other than η, such as functions commonly used in applied statistics: the higher moments about the means, the moment ratios (e.g. K. Pearson's b_1 and b_2), the covariance, the coefficient of correlation, and "Student's" t-statistic. Works on such functions are being done by my university colleagues, and the results will be published shortly.

If ξ is any of the random variables (1), then

$$0 \leq \epsilon\{a(\xi^2 - 1) + b\xi\} = a^2(\alpha_4 - 1) + 2ab\alpha_3 + b^2$$

for all real (a, b). Hence $\alpha_4 - 1 - \alpha_3^2 \geq 0$, and $\alpha_4 - 1 - \alpha_3^2 = 0$ means that there is unit probability that ξ assumes exactly two values. This easily degenerated case is first eliminated in Theorem 3 by the assumption $\alpha_4 - 1 - \alpha_3^2 \neq 0$ and then considered in section 4. In Theorem 4 the condition $\alpha_4 - 1 - \alpha_3^2 \neq 0$ is implied since ξ cannot be a random variable of the nature just described owing to the non-singularity of $P(x)$.

2. Lemmas. Throughout this paper A, B, C, etc. will denote positive numerical constants; A_k, B_k (A_{km}, B_{km}), etc., will denote positive constants depending only on some integer k (integers k and m), and Q_k (Q_{km}) will denote a positive constant depending only on k (k and m) and the distribution function $P(x)$. ϑ, Θ, Θ_k, (Θ_{km}), Λ_k (Λ_{km}) will denote respectively quantities such that $|\vartheta| \leq 1$, $|\Theta| \leq A$, $|\Theta_k| \leq A_k$ ($|\Theta_{km}| \leq A_{km}$), $|\Lambda_k| \leq Q_k$ ($|\Lambda_{km}| \leq Q_{km}$). These symbols do not necessarily stand for the same quantity at each occurrence. Thus $2\vartheta = \Theta$, $k\Theta_k = \Theta_k$ etc. In particular any positive functions of $k, \alpha_3, \cdots, \alpha_k$ is a Q_k.

1.1. Cramér obtains the asymptotic expansion of the characteristic function of the distribution of $\sqrt{n}\xi$, viz. $\epsilon(e^{it\sqrt{n}\xi})$, when (1) do not have the same distribution, valid for $|t| \leq Q_k n^{1/6}$. Since we assume a common distribution for (1), so that the characteristic function is $\left\{p\left(\frac{t}{\sqrt{n}}\right)\right\}^n$, we are able to derive an asymptotic expansion valid for $|t| \leq Q_k\sqrt{n}$. The extension to $\left\{p\left(\frac{t_1}{\sqrt{n}},\right.\right.$

$\cdots, \frac{t_m}{\sqrt{n}}\Big)\Big\}^n$ presents no difficulty. This is done in the following three lemmas, of which Lemma 3 contains the final result.

LEMMA 1.

(17) $$\log p(t) = \sum_{r=2}^{k-1} \frac{\gamma_r(it)^r}{r!} + \Theta_k \beta_k |t|^k, \quad \text{for } |t| \leq \beta_k^{1/k}.$$

PROOF: Since $p(t) = 1 + \sum_{r=1}^{k-1} \frac{\alpha_r(it)^r}{r!} + \frac{\vartheta\beta_k |t|^k}{k!} = 1 + q(t)$ say, we have, for $\beta_k^{1/k}|t| \leq 1$,

$$q(t) \leq \sum_{r=2}^{k} \frac{\beta_r |t|^r}{r!} \leq \sum_{r=2}^{k} \frac{(\beta_k^{1/k}|t|)^r}{r!} < \sum_{r=2}^{\infty} \frac{1}{r!} = e - 2 < \frac{3}{4}.$$

Hence

(18) $$\log p(t) = \sum_{1 \leq j \leq [\frac{1}{2}(k-1)]} (-1)^{j+1} \frac{\{q(t)\}^j}{j} + \Theta |q(t)|^{[\frac{1}{2}(k+1)]}.$$

For $1 \leq j \leq [\frac{1}{2}(k-1)]$ let us expand each $(-1)^{j+1}j^{-1}\{q(t)\}^j$ to get a polynomial $q_j(t)$ of degree $k - 1$ and a remainder $r_j(t)$. In doing this we regard $q(t)$ formally as a polynomial of degree k in t. For this polynomial we have the majorating relation

$$q(t) \ll e^{\beta_k^{1/k}|t|},$$

whence

$$\frac{(-1)^j}{j} \{q(t)\}^j \ll e^{j\beta_k^{1/k}|t|},$$

which gives

(19) $$|r_j(t)| \leq \sum_{r=k}^{\infty} \frac{j^r \beta_k^{r/k} |t|^r}{r!} \leq j^k \beta_k |t|^k e^{j\beta_k^{1/k}|t|} \leq j^k e^j \beta_k |t|^k \leq A_k \beta_k |t|^k.$$

Similarly,

(20) $$|q(t)|^{[\frac{1}{2}(k+1)]} \leq A_k \beta_k |t|^k.$$

From (18), (19), (20) we obtain

(21) $$\log p(t) = \sum_{1 \leq j \leq [\frac{1}{2}(k-1)]} q_j(t) + \Theta_k \beta_k |t|^k.$$

Since the sum in (21) must equal the sum in (17), the Lemma is proved.

LEMMA 2. *Let* $(\zeta_1, \zeta_2, \cdots, \zeta_m)$ *be a random point with* $\epsilon(\zeta_i) = 0$ *and* $\epsilon(|\zeta_i|^k) = \beta_{ki} < \infty$ *for some integer* $k \geq 3$ $(i = 1, \cdots, m)$. *Let* $p(t_1, \cdots, t_m)$ *be the characteristic function. Then for* $|t_i| \leq m^{-2+1/k}\beta_{ki}^{-1/k}\sqrt{n}$ $(i = 1, \cdots, m)$ *we have*

(22) $$n \log p\left(\frac{t_1}{\sqrt{n}}, \cdots, \frac{t_m}{\sqrt{n}}\right) = \sum_{r=2}^{k-1} \frac{i^r U_r}{r!\, n^{\frac{1}{2}(r-2)}} + \frac{\Theta_k V_k}{n^{\frac{1}{2}(k-2)}}$$

where U_r *and* V_r *are the* rth *semi-invariant and the absolute moment respectively of* $\Sigma\, t_i\zeta_i$

PROOF: If $|t_i| \leq m^{-2+1/k}\beta_{ki}^{-1/k}\sqrt{n}$, then $V_k^{1/k} \leq m^{(k-1)/k}(\Sigma\beta_{ki}\,|\,t_i\,|^k)^{1/k} \leq m^{(k-1)/k}(\Sigma\beta_{ki}^{1/k}\,|\,t_i\,|) \leq \sqrt{n}$. Since $p\left(\frac{t_1}{\sqrt{n}}, \cdots, \frac{t_m}{\sqrt{n}}\right)$ is the value at $t = \frac{1}{\sqrt{n}}$ of the characteristic function of $\Sigma t_i\zeta_i$, it follows from Lemma 1 that for $\sqrt{n} \geq V_k^{1/k}$ we have (22).

LEMMA 3. *Let* $(\zeta_1, \cdots, \zeta_m)$ *be a random point with* $\epsilon(\zeta_i) = 0$, $\epsilon(\zeta_i^2) = 1$ *and* $\epsilon(|\,\zeta_i\,|^k) = \beta_{ki} < \infty$ *for some integer* $k \geq 3$. *Let* $\rho_{ij} = \epsilon(\zeta_i\zeta_j)(\rho_{ii} = 1; i, j = 1, \cdots, m)$ *and the matrix* $||\,\rho_{ij}\,||$ *be positive definite. Let*

$$\Delta = \text{det.}\,|\,\rho_{ij}\,|, \qquad \varphi(t_1, \cdots, t_m) = e^{-\frac{1}{2}\sum_{i,j=1}^{m}\rho_{ij}t_it_j}. \tag{23}$$

Let $p(t_1, \cdots, t_m)$ *be the characteristic function. Then there exists a* B_{km} *such that for* $|\,t_i\,| \leq \frac{B_{km}\Delta\sqrt{n}}{\beta_{ki}^{3/k}}$ $(i = 1, \cdots, m)$ *we have*

$$\begin{aligned}\left\{p\left(\frac{t_1}{\sqrt{n}}, \cdots, \frac{t_m}{\sqrt{n}}\right)\right\}^n &= \varphi(t_1, \cdots, t_m)\{1 + \psi(it_1, \cdots, it_m)\} \\ &+ \frac{\Theta_{km}}{n^{\frac{1}{2}(k-2)}}\left\{\sum_{i=1}^{m}\beta_{ki}^{3(k-2)/k}(\,|\,t_i\,|^k\right. \\ &\left.+\,|\,t_i\,|^{k+1} + \cdots + |\,t_i\,|^{3(k-2)})\right\}e^{-\Delta/4m^{m-1}\sum_{i=1}^{m}t_i^2}\end{aligned} \tag{24}$$

where $\psi\,(it_1, \cdots, it_m)$ *is a polynomial each of whose terms has the form*

$$\frac{1}{n^{\nu/2}}\,a_{\nu_1\cdots\nu_m}(it_1)^{\nu_1}\cdots(it_m)^{\nu_m},$$

with $1 \leq \nu \leq k-3$, $3 \leq \nu_1 + \cdots + \nu_m \leq 3(k-3)$, *and* $a_{\nu_1\cdots\nu_m}$ *depending only on* k *and the moments* $\epsilon(\zeta_1^{\mu_1}\cdots\zeta_m^{\mu_m})$, $3 \leq \mu_1 + \cdots + \mu_m \leq k-1$. *If* $k = 3$, *then* $\psi = 0$.

PROOF. If $|\,t_i\,| \leq m^{-2+(1/k)}\beta_{ki}^{-3/k}\Delta\sqrt{n}$, then $|\,t_i\,| \leq m^{-2+(1/k)}\beta_{ki}^{-1/k}\sqrt{n}$ since $\Delta \leq 1$ and $\beta_{ki} \geq 1$. It follows from Lemma 2 and the fact $U_2 = \Sigma\rho_{ij}t_it_j$ that

$$\begin{aligned}\left\{p\left(\frac{t_1}{\sqrt{n}}, \cdots, \frac{t_m}{\sqrt{n}}\right)\right\}^n &= \varphi(t_1, \cdots, t_m)e^s \\ &= \varphi(t_1, \cdots, t_m)\left\{1 + \sum_{j=1}^{k-3}\frac{s^j}{j\,!} + \frac{\vartheta\,|\,s\,|^{k-2}e^{|s|}}{(k-2)\,!}\right\}\end{aligned} \tag{25}$$

where

$$s = \frac{i^3}{\sqrt{n}}\sum_{r=0}^{k-3}\frac{i^r\,U_{r+3}}{(r+3)\,!\,n^{r/2}} + \frac{\Theta_k V_k}{n^{\frac{1}{2}(k-2)}}. \tag{26}$$

Regarding s formally as a polynomial in $n^{-\frac{1}{2}}$ let us expand each $(j!)^{-1}s^j$ $(1 \leq j \leq k-3)$ to get a polynomial s_j of degree $k-3$ in $n^{-\frac{1}{2}}$ and a remainder r_j. For the formal polynomial s we have the majorating relation

$$(27) \qquad s \ll \frac{A_k}{\sqrt{n}} \sum_{r=0}^{k-3} \frac{V_{r+3}}{r!\,n^{r/2}} \ll \frac{A_k}{\sqrt{n}} \sum_{r=0}^{k-3} \frac{V_k^{(r+3)/k}}{r!\,n^{r/2}} \ll \frac{A_k V_k^{3/k}}{\sqrt{n}} e^{V_k^{1/k} n^{-\frac{1}{2}}},$$

whence

$$\frac{1}{j!} s^j \ll A_k \frac{V_k^{3j/k}}{n^{j/2}} e^{jV_k^{1/k} n^{-\frac{1}{2}}},$$

which gives

$$|r_j| \leq \frac{A_k V_k^{3j/k}}{n^{j/2}} \sum_{\nu=k-2-j}^{\infty} \frac{j^\nu V_k^{\nu/k}}{\nu!\,n^{\nu/2}} \leq \frac{A_k V_k^{(k-2+2j)/k}}{n^{\frac{1}{2}(k-2)}} e^{j(V_k^{1/k}/\sqrt{n})}.$$

Since $V_k^{1/k} n^{-\frac{1}{2}} \leq 1$ as shown in the proof of Lemma 2, we have

$$|r_j| \leq \frac{A_k V_k^{(k-2+2j)/k}}{n^{\frac{1}{2}(k-2)}} \leq \frac{A_{km}(\sum_i \beta_{ki} |t_i|^k)^{(k-2+2j)/k}}{n^{\frac{1}{2}(k-2)}}$$

$$\leq \frac{A_{km}(\sum_i \beta_{ki}^{1/k} |t_i|)^{k-2+2j}}{n^{\frac{1}{2}(k-2)}} \leq \frac{A_{km} \sum \beta_{ki}^{(k-2+2j)/k} |t_i|^{k-2+2j}}{n^{\frac{1}{2}(k-2)}}.$$

Since $\beta_{ki} \geq 1$ we have $\beta_{ki}^{(k-2+2j)/k} \leq \beta_{ki}^{3(k-2)/k}$. Hence

$$(28) \qquad |r_j| \leq \frac{A_{km} \sum_i \beta_{ki}^{3(k-2)/k} |t_i|^{k-2+2j}}{n^{\frac{1}{2}(k-2)}}.$$

Similarly

$$(29) \qquad \frac{|s|^{k-2}}{(k-2)!} \leq \frac{A_{km} \sum_i \beta_{ki}^{3(k-2)/k} |t_i|^{3(k-2)}}{n^{\frac{1}{2}(k-2)}}.$$

From (25), (28), (29) we get

$$\left\{p\left(\frac{t_1}{\sqrt{n}}, \cdots, \frac{t_m}{\sqrt{n}}\right)\right\}^n = \varphi(t_1, \cdots, t_m)\left\{1 + \sum_{j=1}^{k-3} s_j + \sum_{j=1}^{k-3} r_j + \frac{\vartheta |s|^{k-2}}{(k-2)!} e^{|s|}\right\}$$

$$= \varphi(t_1, \cdots, t_m)\,\{1 + \psi(it_1, \cdots, it_m)\}$$

$$+ \frac{\Theta_{km}}{n^{\frac{1}{2}(k-2)}} \{\Sigma \beta_{ki}^{3(k-2)/k}(|t_i|^k + |t_i|^{k+1} + \cdots + |t_i|^{3(k-2)})\}\, \varphi(t_1, \cdots, t_m) e^{|s|}$$

where $\psi(it_1, \cdots, it_m)$ stands for Σs_j. The assertion about $\psi(it_1, \cdots, it_m)$ announced in the lemma can now be seen without difficulty. It remains to show that with suitable B_{km} in the lemma, we have

$$\varphi(t_1, \cdots, t_m) e^{|s|} \leq e^{-\Delta/4m^{m-1} \sum_{i=1}^{m} t_i^2}$$

i.e.

$$(30)\qquad -\frac{1}{2}\sum_{i,j=1}^{m}\rho_{ij}t_it_j + |s| \le -\frac{\Delta}{4m^{m-1}}\sum_{i=1}^{m}t_i^2 .$$

From (27) we have

$$(31)\qquad \begin{aligned} |s| &\le \frac{A_k}{\sqrt{n}}V_k^{3/k} \le \frac{A_{km}}{\sqrt{n}}\left(\sum_i \beta_{ki}|t_i|^k\right)^{3/k} \\ &\le \frac{A_{km}}{\sqrt{n}}\left(\sum_i \beta_{ki}^{1/k}|t_i|\right)^3 \le \frac{A_{km}}{\sqrt{n}}\sum_i \beta_{ki}^{3/k}|t_i|^3. \end{aligned}$$

If we choose $B_{km} \le (4m^{m-1}A_{km})^{-1}$ (and $B_{km} \le m^{-2+(1/k)}$ in order that the earlier results may not be affected), the A_{km} here coinciding with the last written A_{km} in (31), we have, for $|t_i| \le B_{km}\beta_{ki}^{-3/k}\Delta\sqrt{n}$,

$$(32)\qquad |s| \le \frac{\Delta}{4m^{m-i}}\sum_{i=1}^{m}t_i^2 .$$

On the other hand, if $\lambda_1, \lambda_2, \cdots, \lambda_m$ are the latent roots of $\|\rho_{ij}\|$ then each $\lambda_i \le m$ since their sum is m. Letting λ_1 be the smallest one we have

$$(33)\qquad \frac{1}{2}\sum_{i,j}\rho_{ij}t_it_j \ge \frac{1}{2}\lambda_1\sum t_i^2 = \frac{\lambda_1\lambda_2\cdots\lambda_m}{2\lambda_2\cdots\lambda_m}\sum t_i^2 \ge \frac{\Delta}{2m^{m-1}}\sum t_i^2 .$$

(32) and (33) imply (30). Hence the lemma is proved.

Let us write down the particular cases $m = 1$ and $m = 2$ of (24):

$$(34)\qquad \begin{aligned} \left\{p\left(\frac{t}{\sqrt{n}}\right)\right\}^n &= e^{-\frac{1}{2}t^2}(1+\psi(it)) \\ &+ \frac{\Theta_k}{n^{\frac{1}{2}(k-2)}}\beta_k^{3(k-2)/k}\{|t|^k + |t|^{k+1} + \cdots + |t|^{3(k-2)}\}e^{-t^2/4}, \left(|t| \le \frac{A_k\sqrt{n}}{\beta_k^{3/k}}\right) \end{aligned}$$

$$(35)\qquad \begin{aligned} \left\{p\left(\frac{t_1}{\sqrt{n}}, \frac{t_2}{\sqrt{n}}\right)\right\}^n &= e^{-\frac{1}{2}(t_1^2+t_2^2+2\rho t_1t_2)}\{1+\psi(it_1, it_2)\} \\ &+ \frac{\Theta_k}{n^{\frac{1}{2}(k-2)}}\left\{\sum_{i=1}^{2}\beta_{ki}^{3(k-2)/k}(|t_i|^k + |t_i|^{k+1} + \cdots + |t_i|^{3(k-2)})\right\}e^{-(1-\rho^2)(t_1^2+t_2^2)/8} \\ &\left(|t_i| \le \frac{A_k(1-\rho^2)\sqrt{n}}{\beta_{ki}^{3/k}}, \quad \rho = \epsilon(\zeta_1\zeta_2)\right). \end{aligned}$$

More specially let us rewrite (34) and (35) with $k = 3$:

$$(36)\qquad \left\{p\left(\frac{t}{\sqrt{n}}\right)\right\}^n = e^{-\frac{1}{2}t^2} + \frac{\Theta}{\sqrt{n}}\beta_3|t|^3e^{-\frac{1}{2}t^2}, \qquad \left(|t| \le \frac{A\sqrt{n}}{\beta_3}\right);$$

$$(37)\qquad \begin{aligned} \left\{p\left(\frac{t_1}{\sqrt{n}}, \frac{t_2}{\sqrt{n}}\right)\right\}^n &= e^{-\frac{1}{2}(t_1^2+t_2^2+2\rho t_1t_2)} \\ &+ \frac{\Theta}{\sqrt{n}}(\beta_{31}|t_1|^3 + \beta_{32}|t_2|^3)e^{-(1-\rho^2)(t_1^2+t_2^2)/8}, \qquad \left(|t_i| \le \frac{A(1-\rho^2)\sqrt{n}}{\beta_{3i}}\right). \end{aligned}$$

In this paper only these last four formulae are needed; they are used in the proofs of Theorems 2, 1, 3, 4 respectively. Cases of $m > 2$ of (24) will be needed for the works on other functions alluded to in the introduction.

1.2. In the following group of lemmas, which culminate in Lemma 7, one finds a generalization of the Riemann-Lebesgue theorem, viz. Lemma 6.

LEMMA 4. *Let $f(x)$ be a polynomial of degree $m > 0$, with real coefficients:*

$$f(x) = \sum_{i=0}^{m} a_i x^{m-i} \qquad (a_0 \neq 0) \tag{38}$$

Then

$$\left| \int_0^1 e^{if(x)}\, dx \right| \leq \frac{A_m}{|a_0|^{1/m}}. \tag{38}$$

PROOF: It is sufficient to prove the inequality for $\int_0^1 \cos f(x)\, dx$. Divide the interval into A_m sub-intervals in each of whose interior none of the derivatives $f^{(i)}(x)$ $(i = 1, \cdots, m)$ vanishes. It is sufficient to consider one of these sub-intervals, say (a, b). Consequently each of the polynomials $f^{(i)}(x)$ are monotonic in (a, b). Let

$$I = \int_a^b \cos f(x)\, dx. \tag{39}$$

Suppose first that $f'(x)$ is positive and increasing for $a < x \leq b$. Then

$$|I| \leq \epsilon + \left| \int_{a+\epsilon}^{b} \frac{f'(x) \cos f(x)\, dx}{f'(x)} \right|$$

$$= \epsilon + \frac{1}{f'(a+\epsilon)} \left| \int_{a+\epsilon}^{b_1} f'(x) \cos f(x)\, dx \right|, \quad (a + \epsilon \leq b_1 \leq b),$$

by the second mean-value theorem. Hence

$$|I| \leq \epsilon + \frac{2}{f'(a + \epsilon)}. \tag{40}$$

Now $0 < f'(a + \frac{1}{2}\epsilon) = f'(a + \epsilon) - \epsilon f''(a + \theta\epsilon)/2$, $\frac{1}{2} \leq \theta \leq 1$. Hence $f'(a + \epsilon) > \frac{1}{2}\epsilon f''(a + \theta\epsilon)$. Since $f''(x)$ is monotonic, we have either $f'(a + \epsilon) > \frac{1}{2}\epsilon f''(a + \epsilon)$ or $f'(a + \epsilon) > \frac{1}{2}\epsilon f''(a + \frac{1}{2}\epsilon)$. In other words, there exists a constant C_2, independent of a or ϵ, such that $\frac{1}{2} \leq C_2 \leq 1$ and $f'(a + \epsilon) > \frac{1}{2}\epsilon f(a + C_2\epsilon)$.

If $f'''(x) \geq 0$, we have, as before $f''(a + C_2\epsilon) > \frac{1}{2}C_2\epsilon f'''(a + C_3\epsilon)$, where C_3 is independent of a or ϵ and $\frac{1}{4} \leq C_3 \leq 1$. If $f'''(x) < 0$, then, since $0 < f''(a + 2C_2\epsilon) = f''(a + C_2\epsilon) + C_2\epsilon f'''(a + \theta_1 C_2\epsilon)$, $\frac{1}{2} \leq \theta_1 \leq 1$, we have $f''(a + C_2\epsilon) > -C_2\epsilon f'''(a + 2\theta_1 C_2\epsilon)$. As $f'''(x)$ is monotonic, either $f''(a + C_2\epsilon) > -C_2\epsilon f'''(a + C_2\epsilon)$ or $f''(a + C_2\epsilon) > -C_2\epsilon f'''(a + 2C_2\epsilon)$. In all cases we obtain $f''(a + C_2\epsilon) > B_3\epsilon\, |f'''(a + C_3\epsilon)|$, where B_3 and C_3 are independent of a or ϵ, and $\frac{1}{4} \leq C_3 \leq 2$. Hence $f'(a + \epsilon) > \frac{1}{2}B_3\epsilon^2\, |f'''(a + C_3\epsilon)|$. Arguing with $\pm f'''(a + C_3\epsilon)$ as we did with $f''(a + C_2\epsilon)$, and so on until we come to $f^{(m)}$,

we obtain $f'(a + \epsilon) > B_m\epsilon^{m-1} | f^{(m)}(a + C_m\epsilon) | = B_m\epsilon^{m-1} | a_0 |$. Substituting in (40) and putting $\epsilon = | a_0 |^{-1/m}$ we obtain $| I | \leq A_m | a_0 |^{-1/m}$. The proof presupposes that $C_m\epsilon < b - a$. If the reverse inequality is true, then $| I | \leq b - a < C_m | a_0 |^{-1/m}$. Hence the lemma is true for $f'(x)$ positive and increasing in (a, b).

If $f'(x)$ is positive and decreasing in (a, b), then $I = \int_0^{b-a} \cos(-f(b - y))\, dy$, $-f(b - y)$ being a polynomial with the leading coefficient $\pm a_0$ and the first derivative $f'(b - y)$, which is positive and increasing. This case reduces therefore to the preceding one. Finally, if $f'(x)$ is negative, we have only to notice that $I = \int_a^b \cos(-f(x))\, dx$. Hence the lemma is proved.

LEMMA 5. *Let $f(x)$ be the polynomial* (38a), *and let $a_r \neq 0$ for some r, $0 \leq r < m$. Then*

$$\left| \int_0^1 e^{if(x)}\, dx \right| \leq \frac{A_m}{| a_r |^{A_m}}. \tag{41}$$

PROOF: We may assume that $| a_r | \geq 1$, (41) being trivial if $| a_r | < 1$. If $r = 0$ this reduces to Lemma 4. Suppose that the lemma is true for $a_0, a_1, \cdots, a_{r-1}$. Let $f_1(x) = a_0x^m + \cdots + a_{r-1}x^{m-r+1}$, $f_2(x) = f(x) - f_1(x)$ and divide (0, 1) into A_m sub-intervals in each of which $f_1(x)$ is monotonic. It is sufficient to consider one of these sub-intervals, say, (a, b). We have

$$I = \int_a^b \cos\{f_1(x) + f_2(x)\}\, dx$$

$$= \int_a^b \cos f_1(x) \cos f_2(x)\, dx - \int_a^b \sin f_1(x) \sin f_2(x)\, dx.$$

We have only to consider the integral of cosines, say J. Divide (a, b) into sub-intervals in each of whose interior $\cos f_1(x)$ is monotonic and does not vanish. The number of such intervals does not exceed $(\frac{1}{2}\pi)^{-1} | f_1(b) - f_1(a) | \leq (\frac{1}{2}\pi)^{-1}(| f_1(b) | + | f_1(a) |) < 2(| a_0 | + \cdots + | a_{r-1} |)$. Then, by the second mean-value theorem,

$$| J | \leq 2(| a_0 | + \cdots + | a_{r-1} |) \left| \int_a^{b_1} \cos f_2(x)\, dx \right| \qquad (a \leq b_1 \leq b).$$

Hence, applying Lemma 4 to $f_2(x)$, we get

$$| I | \leq \frac{A_m(| a_0 | + \cdots + | a_{r-1} |)}{| a_r |^{1/(m-r)}} \leq \frac{A_m(| a_0 | + \cdots + | a_{r-1} |)}{| a_r |^{1/m}}. \tag{42}$$

On the hypothesis of induction we have $| I | \leq A_m | a_i |^{-B_m}$ $(i = 0, \cdots, r - 1)$. If $| a_i | \geq | a_r |^{1/2m}$ for some $i < r$, then $| I | \leq A_m | a_r |^{-B_m/2m}$; if $| a_i | < | a_r |^{1/2m}$, then by (42), $| I | \leq A_m | a_r |^{-1/2m}$. The proof is therefore complete.

LEMMA 6. *Let $f(x)$ be the polynomial* (38a) *and $g(x)$ be summable over* $(-\infty, \infty)$. *Then for every r we have*

$$(43)\qquad \lim_{|a_r|\to\infty} \int_{-\infty}^{\infty} e^{if(x)} g(x)\,dx = 0, \quad \textit{uniformly in } a_i (i \neq r).$$

PROOF: By Lemma 5 We have

$$\lim_{|a_r|\to\infty} \int_0^1 e^{if(x)}\,dx = 0, \quad \text{uniformly in } a_i (i \neq r).$$

Hence

$$(44)\qquad \lim_{|a_r|\to\infty} \int_a^b e^{if(x)}\,dx = 0, \quad \text{uniformly in } a_i (i \neq r)$$

for if $a \neq 0$ and $b \neq 0$, then (a, b) is the sum or the difference of two intervals of the form $(0, c)$ or $(c, 0)$, and for the latter intervals the transformation $x = \pm cy$ reduces the interval of integration to $(0, 1)$.

Let G be any open set of finite measure. Then G is the sum of a sequence $\{I_\nu\}$ of non-overlapping intervals. Since $\Sigma mI_\nu = mG < \infty$, we have

$$\sum_{\nu \geq n} mI_\nu < \epsilon, \qquad n \geq N.$$

Hence

$$\left| \int_G e^{if(x)}\,dx \right| < \epsilon + \sum_{\nu=1}^{N} \left| \int_{I_\nu} e^{if(x)}\,dx \right|$$

which, together with (44), implies

$$(45)\qquad \lim_{|a_r|\to\infty} \int_G e^{if(x)}\,dx = 0 \quad \text{uniformly in } a_i (i \neq r).$$

Let S be any set of finite measure. Then there is an open set G such that $G \supset S$ and $m(G - S) < \epsilon$. Hence

$$\left| \int_S e^{if(x)}\,dx \right| < \epsilon + \left| \int_G e^{if(x)}\,dx \right|.$$

Hence, by (45),

$$(46)\qquad \lim_{|a_r|\to\infty} \int_S e^{if(x)}\,dx = 0 \quad \text{uniformly in } a_i (i \neq r).$$

Now let $h(x)$ be any positive "simple" summable function, i.e. $h(x) = a_\nu > 0$ for $x \in S$ $(\nu = 1, 2, \cdots, n)$ and $h(x) = 0$ otherwise. Since $h(x)$ is summable, each S_ν must be of finite measure. Hence

$$\left| \int_{-\infty}^{\infty} e^{if(x)} h(x)\,dx \right| \leq \sum_{\nu=1}^{n} a_\nu \left| \int_{S_\nu} e^{if(x)}\,dx \right|$$

which, together with (46), implies

$$\lim_{|a_r|\to\infty} \int_{-\infty}^{\infty} e^{if(x)} h(x)\,dx = 0 \quad \text{uniformly in } a_i (i \neq r).$$

Finally, let $g(x)$ be any summable function ≥ 0. Then by a well-known theorem[6] we have $g(x) = \lim h_n(x)$, where $\{h_n(x)\}$ is an ascending sequence of positive summable simple functions. Hence

$$\left|\int_{-\infty}^{\infty} e^{if(x)} g(x)\, dx\right| \leq \left|\int_{-\infty}^{\infty} e^{if(x)} h_n(x)\, dx\right| + \int_{-\infty}^{\infty} (g(x) - h_n(x))\, dx.$$

By monotonic convergence the last integral tends to 0 as $n \to \infty$. Hence

$$\left|\int_{-\infty}^{\infty} e^{if(x)} g(x)\, dx\right| \leq \epsilon + \left|\int_{-\infty}^{\infty} e^{if(x)} h_n(x)\, dx\right|,$$

which implies (43). If $g(x)$ is any summable function, we have only to consider the customary expression of $g(x)$ as the difference of two non-negative functions. This completes the proof.

LEMMA 7. *Let $P(x)$ be a non-singular distribution function of a random variable X, and let*

$$(47) \qquad p(t_1, t_2, \cdots, t_m) = \int_{-\infty}^{\infty} e^{i \sum_{r=1}^{m} t_r x^r}\, dP.$$

Then for every r and every positive constant c we have

$$(48) \qquad \underset{|t_r| \geq c}{\text{l.u.b.}} |p(t_1, \cdots, t_m)| < 1.$$

PROOF: We have $P(x) = a_1 P_1(x) + a_2 P_2(x)$, where $P_1(x)$ is absolutely continuous, P_2 is singular, $a_1 > 0$, $a_1 + a_2 = 1$. Hence

$$|p(t_1, t_2, \cdots, t_m)| \leq a_1 \left|\int_{-\infty}^{\infty} e^{i \sum_{r=1}^{m} t_r x^r} P_1'(x)\, dx\right| + a_2.$$

By Lemma 6 we may find $C > 0$ such that

$$|p(t_1, t_2, \cdots, t_m)| \leq \tfrac{1}{2} a_1 + a_2 < 1, \quad \text{if any} \quad |t_i| > C.$$

Suppose that

$$\underset{|t_r| \geq c}{\text{l.u.b.}}\, p(t_1, \cdots, t_m) = 1,$$

then $c < C$ and we must have

$$(49) \qquad \underset{c \leq |t_r| \leq C, |t_i| \leq C (i \neq r)}{\text{l.u.b.}} |p(t_1, \cdots, t_m)| = 1.$$

Since $p(t_1, \cdots, t_m)$ is a continuous function, it must attain its least upper bound in any bounded closed set. It follows that there is a point $(t_1^0, \cdots, t_m^0)$ such that[7] $t_r^0 \neq 0$ ($|t_r^0| \geq c$) and $p(t_1^0, \cdots, t_m^0) = 1$. But this implies that the distribution of $\Sigma t_i^0 X^i$ is discrete, i.e. that the distribution of X itself is discrete,

[6] H. Kestelman: *Modern Theories of Integration* (1937), p. 108.

[7] Cf. (C), p. 26.

which contradicts the non-singularity of $P(x)$. Hence (49) is false and (48) is true.

1.3. In his cited work Berry[8] shows that if $F(x)$ is any distribution function and if $\Phi(x)$ is the function (6), then there is a constant a such that

$$\left| \int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2} \{F(x + a) - \Phi(x + a)\} \, dx \right| \geq \sqrt{\frac{2}{\pi}} \, T\delta \left\{ 3 \int_0^{P\delta} \frac{1 - \cos x}{x^2} \, dx - \pi \right\} \tag{50}$$

where $\delta = \sqrt{\frac{\pi}{2}} \, \text{l.u.b.} \, | F(x) - \Phi(x) |$. This is easily extended to the following lemma, which needs no further proof.

Lemma 8. *Let $F(x)$ be a distribution function and $F_1(x)$ be a function having the following properties:* (i) *$F_1(x)$ is bounded for all x,* (ii) *$F_1(x) \to 1$ as $x \to \infty$, $F_1(x) \to 0$ as $x \to -\infty$,* (iii) *$F_1(x)$ has a bounded derivative, $| F_1'(x) | \leq M$. Let*

$$\delta = \frac{1}{2M} \, \text{l.u.b.} \, | F(x) - F_1(x) | .$$

Then there exists a constant a such that

$$\left| \int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2} \{F(x + a) - F_1(x + a)\} \, dx \right| \geq 2MT\delta \left\{ 3 \int_0^{T\delta} \frac{1 - \cos x}{x^2} \, dx - \pi \right\}. \tag{51}$$

1.4. In section 3 we define, for given ϵ, k, λ and z, a function

$$G(x, y) = e^{-\epsilon y^{2k}} \quad \text{if} \quad z < x \leq z + \lambda y^2, \quad G(x, y) = 0 \quad \text{otherwise.} \tag{52}$$

The introduction of $G(x, y)$ and the appraisal of its Fourier transform constitute the essence of our method of solving the problem of the asymptotic expansion of the distribution function $G(x)$. The solution of the same problem about other functions of (1) alluded to in section 3 is based on the introduction of functions playing the role of $G(x, y)$. We now prove the following lemma:

Lemma 9. *Let $G(x, y)$ be defined by* (52) *and let*

$$g(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-it_1 x - it_2 y} G(x, y) \, dx \, dy. \tag{53}$$

Then

(i) $$| g(t_1, t_2) | \leq \frac{\lambda A_k}{\epsilon^{3/2k}}$$

(ii) $$| g(t_1, t_2) | \leq \frac{A}{| t_2 |^3} \left(\lambda + \frac{\lambda^2 | t_1 |}{\epsilon^{1/3}} + \frac{\lambda^3 | t_1 |^2}{\epsilon^{2/3}} \right) \quad \text{if } k = 3,$$

(iii) $$| g(t_1, t_2) | \leq \frac{A_k}{| t_2 |^2} \left(\frac{\lambda}{\epsilon^{1/2k}} + \frac{\lambda^2 | t_1 |}{\epsilon^{3/2k}} \right).$$

[8] (B), p. 128.

PROOF:

(i) $$|g(t_1, t_2)| \leq \int_{R_2} G(x, y)\, dx\, dy = \lambda \int_{-\infty}^{\infty} y^2 e^{-\epsilon y^{2k}}\, dy = \frac{A_k \lambda}{\epsilon^{3/2k}}$$

(ii) Putting $k = 3$ we have

$$g(t_1, t_2) = \frac{e^{-it_1 x}}{it_1} \int_{-\infty}^{\infty} e^{-\epsilon y^6 - it_2 y}(1 - e^{-it_1 \lambda y^2})\, dy,$$

$$|g(t_1, t_2)| \leq \frac{1}{|t_1|\,|t_2|^3} \left| \int_{-\infty}^{\infty} u(y) v'''(y)\, dy \right|,$$

where $u(y) = e^{-\epsilon y^6}(1 - e^{-it_1\lambda y^2})$, $v(y) = e^{-it_2 y}$. On integrating by parts we obtain

$$(54) \quad |g(t_1, t_2)| \leq \frac{1}{|t_1|\,|t_2|^3} \left| \int_{-\infty}^{\infty} v(y) u'''(y)\, dy \right| \leq \frac{1}{|t_1|\,|t_2|^3} \int_{-\infty}^{\infty} |u'''(y)|\, dy.$$

Elementary calculation establishes that

$$\frac{|u'''(y)|}{|t_1|} \leq e^{-\epsilon y^6}(216\lambda\epsilon^3 |y|^{17} + 756\lambda\epsilon^2 |y|^{11}$$
$$+ 336\lambda\epsilon |y|^5 + 8\lambda^3 |t_1|^2 |y|^3 + 12\lambda^2 |t_1|\,|y|).$$

Substituting in (54) and making the transformation $y = \epsilon^{-1/6} x$ we get the result.

(iii) We have

$$|g(t_1, t_2)| \leq \frac{1}{|t_1|} \left| \int_{-\infty}^{\infty} e^{-\epsilon y^{2k} - it_2 y}(1 - e^{-it_1\lambda y^2})\, dy \right|.$$

Integrating by parts twice we obtain

$$|g(t_1, t_2)| \leq \frac{1}{|t_1|\,|t_2|^2} \int_{-\infty}^{\infty} \left| \frac{d^2}{dy^2} \{e^{-\epsilon y^{2k}}(1 - e^{-it_1\lambda y^2})\} \right| dy.$$

By elementary calculations we get

$$|g(t_1, t_2)| \leq \frac{1}{|t_2|^2} \int_{-\infty}^{\infty} (4k^2\lambda\epsilon y^{4k} + 2k(k+3)\lambda\epsilon y^{2k} + 4\lambda^2 |t_1| y^2 + 2\lambda) e^{-\epsilon y^{2k}}\, dy$$

which, on the transformation $y = \epsilon^{-1/2k} x$, gives the result.

1.5. We prove a few additional lemmas used in the proof of Theorems 3 and 4.

LEMMA[9] 10. *Let $u(x_1, \cdots, x_m) \geq 0$ be summable in the m-dimensional space and let*

$$(55) \quad v(t_1, \cdots, t_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-it_1 x_1 - \cdots - it_m x_m} u(x_1, \cdots, x_m)\, dx_1 \cdots dx_m.$$

[9] Although the author believes that this lemma is almost classical, a proof is given owing to lack of reference.

If $v(t_1, \cdots, t_m)$ is summable in the m-dimensional space, then

$$u(x_1, \cdots, x_m) = \frac{1}{(2\pi)^m} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{it_1x_1+\cdots+it_mx_m} v(t_1, \cdots t_m)\, dt_1 \cdots dt_m. \tag{56}$$

PROOF: Except for a constant factor the function $u(x_1, \cdots, x_m)$ may be regarded as a probability density function. Hence by the well-known inversion formula of (55),

$$\begin{aligned} &\underset{a_i \le x_i \le b_i \ (i=1,\cdots,m)}{\int \cdots \int} u(x_1, \cdots, x_m)\, dx_1 \cdots dx_m \\ &\quad = \frac{1}{(2\pi)^m} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{m} \frac{e^{it_jb_j} - e^{it_ja_j}}{it_j} \right) v(t_1, \cdots, t_m)\, dt_1 \cdots dt_m . \end{aligned} \tag{57}$$

Now $u(x_1, \cdots, x_m)$ is almost everywhere the symmetric derivative of the interval function in the left-hand side of (57):

$$u(x_1, \cdots, x_m) = \lim_{\epsilon \to 0} \frac{1}{(2\epsilon)^m} \underset{x_i - \epsilon \le y_i \le x_i + \epsilon \ (i=1,2,\cdots,m)}{\int \cdots \int} u(y_1, \cdots, y_m)\, dy_1 \cdots dy_m .$$

Hence

$$\begin{aligned} u(x_1, \cdots, x_m) = \frac{1}{(2\pi)^m} \lim_{\epsilon \to 0} \frac{1}{(2\epsilon)^m} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \\ \cdot \left(\prod_{j=1}^{m} \frac{e^{it_j\epsilon} - e^{-it_j\epsilon}}{it_j} \right) e^{it_1x_1+\cdots+it_mx_m} v(t_1, \cdots, t_m)\, dt_1 \cdots dt_m . \end{aligned} \tag{58}$$

Owing to dominated convergence the order of the limit sign and the integration sign in (58) may be inverted: Hence (56) is true.

LEMMA 11. *We have*

$$\int_{-\infty}^{\infty} e^{-itu} \frac{1 - \cos Tu}{u^2} du = \begin{cases} \pi(T - |t|) & \text{if } |t| \le T, \\ 0 & \text{if } |t| > T. \end{cases} \tag{59}$$

PROOF: The Fourier transform of the function in the right-hand side of (59) is

$$\pi \int_{-T}^{T} e^{itu}(T - |t|)\, dt = \frac{2\pi}{u^2} (1 - \cos Tu).$$

Hence (59) follows from (56).

LEMMA 12.

$$| \epsilon(\xi_1 + \cdots + \xi_n)^k | \le A_k n^{k/2} \beta_k \tag{60}$$

PROOF. As (60) is true for $k = 1$, let us assume, for induction, that it is true for $1, 2, \cdots, k$. Then, by symmetry,

$$\epsilon(\xi_1 + \cdots + \xi_n)^{k+1} = n\epsilon\{\xi_1(\xi_1 + \cdots + \xi_n)^k\} = n \sum_{r=0}^{k} \binom{k}{r} \epsilon(\xi_1^{r+1} U^{k-r})$$

where $U = \xi_2 + \cdots + \xi_k$. Since $\epsilon(\xi_1) = 0$, we have

$$\epsilon(\xi_1 + \cdots + \xi_n)^{k+1} = n \sum_{r=1}^{k} \binom{k}{r} \epsilon(\xi_1^{r+1} U^{k-r}).$$

On the hypotheses of induction we have $| \epsilon(U^{k-r}) | \leq A_k(n - 1)^{\frac{1}{2}(k-r)}\beta_{k-r} < A_k n^{\frac{1}{2}(k-1)}\beta_{k-r}$. Hence

$$|\epsilon(\xi_1 + \cdots + \xi_n)^{k+1}| \leq k! A_k n^{\frac{1}{2}(k+1)} \Sigma \beta_{r+1}\beta_{k-r} \leq A_{k+1} n^{\frac{1}{2}(k+1)} \beta_{k+1}.$$

Therefore the induction is complete.

3. Elementary Proof of Theorem 1. 2.1 We have defined

$$F(x) = Pr\{\sqrt{n}\xi \leq x\}, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\, dy \tag{61}$$

with the characteristic functions

$$f(t) = \left\{ p\left(\frac{t}{\sqrt{n}}\right) \right\}^n, \qquad \varphi(t) = e^{-\frac{1}{2}t^2}. \tag{62}$$

Following Berry[10] we use the equation

$$\int_{-\infty}^{\infty} \{F(x) - \Phi(x)\} e^{itx}\, dx = \frac{f(t) - \varphi(t)}{-it}. \tag{63}$$

Let $\psi(it)$ be the polynomial in (34), and let us define $\Psi(x)$ as the function obtained from $\psi(it)$ through the replacement of each power $(it)^\nu$ by $(-1)^\nu \Phi^{(\nu)}(x)$. Integration by parts shows $(-1)^{\nu-1} \int_{-\infty}^{\infty} e^{itx}\Phi^{(\nu)}(x)\, dx = (it)^{\nu-1}\varphi(t)$, whence

$$\int_{-\infty}^{\infty} \Psi(x) e^{itx}\, dx = \frac{\psi(it)\varphi(t)}{-it}. \tag{64}$$

From (63) and (64) we obtain

$$\int_{-\infty}^{\infty} \{F(x) - \Phi(x) - \Psi(x)\} e^{itx}\, dx = \frac{f(t) - \varphi(t)\{1 + \psi(it)\}}{-it}. \tag{65}$$

The function $\Psi(x)$ defined here is precisely the $\Psi(x)$ appearing in (5) under Theorem 1. Our task is to prove that

$$| F(x) - \Phi(x) - \Psi(x) | \leq \frac{Q_k}{n^{(k-2)/2}}. \tag{66}$$

Following Berry[11] we replace x by $x + a$ in (65), getting

$$\begin{aligned} &\int_{-\infty}^{\infty} \{F(x + a) - \Phi(x + a) - \Psi(x + a)\} e^{itx}\, dx \\ &\qquad = \frac{e^{-ita}[f(t) - \varphi(t)\{1 + \psi(it)\}]}{-it} \end{aligned} \tag{67}$$

[10] (B), p. 127, Equation (23).
[11] (B), p. 127.

multiply both sides of (67) by $T - |t|$ and integrate with respect to t in $(-T, T)$:

$$2\int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2} \{F(x + a) - \Phi(x + a) - \Psi(x + a)\}\, dx$$

$$= \int_{-T}^{T} \frac{(T - |t|)e^{-ita}[f(t) - \varphi(t)\{1 + \psi(it)\}]}{-it}\, dt$$

the reversion of order of integration involved is obviously justifiable. Hence

$$\left|\int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2} \{F(x + a) - \Phi(x + a) - \Psi(x + a)\}\, dx\right| \tag{68}$$

$$\leq T \int_0^T \frac{|f(t) - \varphi(t)\{1 + \psi(it)\}|}{t}\, dt.$$

2.2. When in particular $k = 3$, (68) becomes

$$\left|\int_{-\infty}^{\infty} \frac{1 - \cos Tx}{x^2} \{F(x + a) - \Phi(x + a)\}\, dx\right| \leq T\int_0^T \frac{|f(t) - \varphi(t)|}{t}\, dt. \tag{69}$$

If we choose a to be the a in (50), the left-hand side of (69) is not less than

$$\sqrt{\frac{2}{\pi}}\, T\delta \left\{3\int_0^{T\delta} \frac{1 - \cos x}{x^2}\, dx - \pi\right\}, \qquad \delta = \sqrt{\frac{\pi}{2}}\, \text{l.u.b.}\,|F(x) - \Phi(x)|.$$

On the other hand, taking $T = \dfrac{A\sqrt{n}}{\beta_3}$ as in (36) the right-hand side of (69) is not greater than

$$A\int_0^{\infty} t^2 e^{-\frac{1}{4}t^2}\, dt = A.$$

Hence

$$T\delta\left\{3\int_0^{T\delta} \frac{1 - \cos x}{x^2}\, dx - \pi\right\} \leq A. \tag{70}$$

Now the left-hand side of (70), as a function of $T\delta$, is positive and increasing for sufficiently large $T\delta$, and becomes infinite as $T\delta \to \infty$. Hence (70) implies that $T\delta \leq A$, i.e.

$$\text{l.u.b.}\,|F(x) - \Phi(x)| \leq \frac{A}{T} = \frac{A\beta_3}{\sqrt{n}},$$

giving Theorem 2.

2.3. Coming back to the general case, we see that the function $\Phi(x) + \Psi(x)$ has a bounded derivative: $|\Phi'(x) + \Psi'(x)| \leq Q_k$, and also has all the properties of the function $F_1(x)$ in Lemma 8. On choosing a in (69) to be the a in (51) we obtain

$$(71)\quad Q_k T\delta\left\{3\int_0^{T\delta}\frac{1-\cos x}{x^2}\,dx-\pi\right\}\le T\int_0^T\frac{|f(t)-\varphi(t)\{1+\psi(it)\}|}{t}\,dt,$$

where

$$\delta = Q_k \text{ l.u.b. } |F(x) - \Phi(x) - \Psi(x)|.$$

Let us take $T = (A_k \beta_k^{-3/k}\sqrt{n})^{k-2}$ with A_k in accordance with (34). Then

$$(72)\quad \begin{aligned} &T\int_0^T\frac{|f(t)-\varphi(t)\{1+\psi(it)\}|}{t}\,dt \\ &\qquad = Q_k n^{\frac{1}{2}(k-2)}\int_0^{T^{1/(k-2)}} + Q_k n^{\frac{1}{2}(k-2)}\int_{Q_k\sqrt{n}}^{T} = J_1 + J_2 \quad \text{say.}\end{aligned}$$

By (34) we have

$$(73)\quad J_1 \le Q_k\int_0^\infty (t^{k-1}+\cdots+t^{3k-7})e^{-\frac{1}{2}t^2}\,dt = Q_k.$$

Also,

$$(74)\quad J_2 \le Q_k n^{\frac{1}{2}(k-2)}\int_{Q_k\sqrt{n}}^{T}\frac{|p(t/\sqrt{n})|^n}{t}\,dt + Q_k n^{\frac{1}{2}(k-2)}\int_{Q_k\sqrt{n}}^{T}\frac{\varphi(t)\,|1+\psi(it)|}{t}\,dt.$$

The second term in the right-hand side of (74) is evidently $\le Q_k$. The first term does not exceed

$$(75)\quad Q_k n^{\frac{1}{2}(k-3)}\,T \operatorname*{l.u.b.}_{t\ge Q_k} |p(t)|^n.$$

At this step we make use of the non-singularity of $P(x)$ and apply Lemma 7 for $m = 1$. We have

$$\operatorname*{l.u.b.}_{t\ge Q_k}|p(t)| = e^{-Q_k}.$$

Hence (75) does not exceed $Q_k n^{\frac{1}{2}(2k-5)}e^{-Q_k n} \le Q_k$. We have therefore

$$(76)\quad T\delta\left\{3\int_0^{T\delta}\frac{1-\cos x}{x^2}\,dx-\pi\right\}\le Q_k,\qquad T = Q_k n^{\frac{1}{2}(k-2)}.$$

Arguing with (76) as we did with (70) we conclude that

$$\text{l.u.b.}\,|F(x)-\Phi(x)-\Psi(x)| \le \frac{Q_k}{T} = \frac{Q_k}{n^{\frac{1}{2}(k-2)}}.$$

(72) is valid for $T \ge 1$. If $T < 1$, we have only to suppress the term J_2. Hence Theorem 1 is proved.

4. Proof of Theorem 3 and Theorem 4. 3.1. In connection with the random variables (1), we assume that $\beta_{2k} < \infty$ for some integer $k \ge 3$ and define

$$(77)\quad \eta = \frac{1}{n}\sum_{r=1}^{n}(\xi_r - \bar{\xi})^2,\qquad G(z) = Pr\left\{\frac{\sqrt{n}(\eta-1)}{\sqrt{\alpha_4 - 1}} \le z\right\}.$$

Now,

$$\eta = \frac{1}{n} \sum \xi_r^2 - \bar{\xi}^2 = 1 + \sqrt{\frac{\alpha_4 - 1}{n}}\, X - \frac{Y^2}{n}$$

where

(78) $$X = \frac{1}{\sqrt{n}} \sum \frac{(\xi_r^2 - 1)}{\sqrt{\alpha_4 - 1}}, \qquad Y = \sqrt{n}\, \bar{\xi}.$$

Hence

(79) $$G(z) = Pr\{X - \lambda Y^2 \le z\}$$

with

(80) $$\lambda = \frac{1}{\sqrt{n(\alpha_4 - 1)}}.$$

Let W be the probability function of the distribution of the random point (X, Y) and $f(t_1, t_2)$ be the characteristic function:

(81) $$W(S) = Pr\{(X, Y) \epsilon S\} \quad \text{for every Borel set } S \text{ in } R_2,$$

(82) $$f(t_1, t_2) = \epsilon(e^{it_1X + it_2Y}) = \left\{p\left(\frac{t_1}{\sqrt{n}}, \frac{t_2}{\sqrt{n}}\right)\right\}^n$$

(83) $$p(t_1, t_2) = \int_{-\infty}^{\infty} e^{it_1(x^2-1)/(\sqrt{\alpha_4 - 1}) + it_2x}\, dP.$$

Let $G_1(z)$ be the distribution function of X. Then

(84) $$G(z) - G_1(z) = \iint\limits_{z < x \le z + \lambda y^2} dW = K(z), \quad \text{say}.$$

Let

(85) $$K_\epsilon(z) = \iint\limits_{z < x \le z + \lambda y^2} e^{-\epsilon y^{2k}}\, dW.$$

If we define (for fixed z) the function $G(x, y)$ by

(86) $$G(x, y) = e^{-\epsilon y^{2k}} \quad \text{if} \quad z < x \le z + \lambda y^2, \qquad G(x, y) = 0 \quad \text{otherwise},$$

then

(87) $$K_\epsilon(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(x, y)\, dW.$$

Letting

(88) $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-it_1x - it_2y} G(x, y)\, dx\, dy = g(t_1, t_2),$$

we replace x by $x - u$ in the integral and get

$$(89) \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-it_1x-it_2y}G(x-u, y)\,dx\,dy = e^{-it_1u}g(t_1, t_2).$$

Multiplying both sides by $\dfrac{1-\cos Tu}{u^2}$ and integrating with respect to u we obtain, with the help of (59), Lemma 11,

$$(90) \qquad \begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-it_1x-it_2y}\,dx\,dy \int_{-\infty}^{\infty} \frac{1-\cos Tu}{u^2} G(x-u, y)\,du \\ &\qquad = \begin{cases} \pi(T - |t_1|)g(t_1, t_2) & \text{if } |t_1| \leq T, \\ 0 & \text{if } |t_1| > T; \end{cases} \end{aligned}$$

the reversion of order of integration in the left-hand side is obviously justifiable. By Lemma 9 the right-hand side of (90) is summable in the whole plane of (t_1, t_2). Hence, by Lemma 10,

$$(91) \qquad \begin{aligned} &\int_{-\infty}^{\infty} \frac{1-\cos Tu}{u^2} G(x-u, y)\,du \\ &\qquad = \frac{1}{4\pi}\iint_{|t_1|\leq T} (T - |t_1|)g(t_1, t_2)e^{it_1x+it_2y}\,dt_1\,dt_2 . \end{aligned}$$

If we integrate both sides with respect to the probability function W, we obtain, on reversing the order of integration,

$$(92) \qquad \begin{aligned} &\int_{-\infty}^{\infty} \frac{1-\cos Tu}{u^2}\,du \iint_{R_2} G(x-u, y)\,dW \\ &\qquad = \frac{1}{4\pi}\iint_{|t_1|\leq T} (T - |t_1|)g(t_1, t_2)f(t_1, t_2)\,dt_1\,dt_2 . \end{aligned}$$

By (86) and (87),

$$(93) \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-u, y)\,dW = K_\epsilon(u+z).$$

Hence

$$(94) \qquad \int_{-\infty}^{\infty} \frac{1-\cos Tu}{u^2} K_\epsilon(u+z)\,du = \frac{1}{4\pi}\iint_{|t_1|\leq T} (T - |t_1|)g(t_1, t_2)f(t_1, t_2)\,dt_1\,dt_2 .$$

We now take the functions

$$(95) \qquad \varphi(t_1, t_2) = e^{-\frac{1}{2}(t_1^2+t_2^2+2\rho t_1t_2)}$$

and $\psi(it_1, it_2)$ as in (35), where

$$\rho = \int_{-\infty}^{\infty} \frac{(x^2 - 1)x}{\sqrt{\alpha_4 - 1}}\, dP = \frac{\alpha_3}{\sqrt{\alpha_4 - 1}}. \tag{96}$$

Since the condition $\alpha_4 - 1 - \alpha_3^2 \neq 0$ is assumed in Theorem 3 and implied in Theorem 4, we have $|\rho| < 1$. Let

$$w(x, y) = \frac{1}{2\pi\sqrt{1 - \rho^2}}\, e^{-(1/2(1-\rho^2))(x^2+y^2-2\rho xy)} \tag{97}$$

and let $\gamma(x, y)$ be the function obtained from $\psi(it_1, it_2)$ through the replacement of each power $(it_1)^{\nu_1}\,(it_2)^{\nu_2}$ by $(-1)^{\nu_1+\nu_2} W_{\nu_1\nu_2}(x, y) = (-1)^{\nu_1+\nu_2} \dfrac{\partial^{\nu_1+\nu_2} w(x, y)}{\partial x^{\nu_1} \partial y^{\nu_2}}$.

Since

$$w(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-it_1x - it_2y} \varphi(t_1, t_2)\, dt_1 dt_2, \tag{98}$$

we have

$$w_{\nu_1\nu_2}(x, y) = \frac{(-1)^{\nu_1+\nu_2}}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (it_1)^{\nu_1} (it_2)^{\nu_2} e^{-it_1x - it_2y} \varphi(t_1, t_2)\, dt_1 dt_2, \tag{99}$$

whence, by Fourier inversion,

$$(it_1)^{\nu_1} (it_2)^{\nu_2} \varphi(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{it_1x + it_2y} w_{\nu_1\nu_2}(x, y)\, dxdy. \tag{100}$$

From the definition of $\gamma(x, y)$ it follows therefore

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{it_1x+it_2y} \{w(x, y) + \gamma(x, y)\}\, dx\, dy = \varphi(t_1, t_2)\{1 + \psi(it_1, it_2)\}. \tag{101}$$

A comparison of (101) with $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{it_1x+it_2y}\, dW = f(t_1, t_2)$ shows that (94) will remain true if $K_\epsilon(u)$ be replaced by

$$\iint\limits_{u < x \leq u + \lambda y^2} e^{-\epsilon y^{2k}} (w(x, y) + \gamma(x, y))\, dx\, dy = L_\epsilon(u), \text{ say}, \tag{102}$$

and $f(t_1, t_2)$ be replaced by $\varphi(t_1, t_2)\{1 + \psi(it_1, it_2)\}$. Hence

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{1 - \cos Tu}{u^2} \{K_\epsilon(u + z) - L_\epsilon(u + z)\}\, du \\ &= \frac{1}{4\pi} \iint\limits_{|t_1| \leq T} (T - |t_1|) g(t_1, t_2) \{f(t_1, t_2) \\ &\qquad - \varphi(t_1, t_2)[1 + \psi(it_1, it_2)]\}\, dt_1 dt_2. \end{aligned} \tag{103}$$

Let also

$$(104)\qquad H(z) = \iint\limits_{x-\lambda y^2 \le z} \{w(x, y) + \gamma(x, y)\}\, dx\, dy,$$

$$H_1(z) = \iint\limits_{x \le z} \{w(x, y) + \gamma(x, y)\}\, dx\, dy,$$

$$(105)\qquad L(z) = H(z) - H_1(z) = \iint\limits_{z < x \le z+\lambda y^2} \{w(x, y) + \gamma(x, y)\}\, dx\, dy.$$

3.2. We now consider the particular case $k = 3$ and prove Theorem 3. For $k = 3$ we have $\psi \equiv \gamma \equiv 0$ and so

$$(106)\qquad \begin{aligned} H(z) &= \iint\limits_{x-\lambda y^2 \le z} w(x, y)\, dx\, dy, \\ H_1(z) &= \iint\limits_{x \le z} w(x, y)\, dx\, dy = \Phi(z), \end{aligned}$$

$$(107)\qquad \begin{aligned} L(z) &= H(z) - H_1(z), \\ L_\epsilon(z) &= \iint\limits_{z < x \le z+\lambda y^2} e^{-\epsilon y^6} w(x, y)\, dx\, dy, \end{aligned}$$

$$(108)\qquad \begin{aligned} &\int_{-\infty}^{\infty} \frac{1 - \cos Tu}{u^2} \{K_\epsilon(u + x) - L_\epsilon(u + x)\}\, du \\ &\quad = \frac{1}{4\pi} \iint\limits_{|t_1| \le T} (T - |t_1|) g(t_1, t_2) \{f(t_1, t_2) - \varphi(t_1, t_2)\} dt_1\, dt_2 . \end{aligned}$$

Now

$$K_\epsilon(u) - L_\epsilon(u) = \{G(u) - \Phi(u)\} - \{H(u) - \Phi(u)\} - \{G_1(u) - \Phi(u)\}$$
$$- \{K(u) - K_\epsilon(u)\} + \{L(u) - L_\epsilon(u)\},$$

$$0 \le H(u) - \Phi(u) = \frac{1}{2\pi\sqrt{1 - \rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\, dy \int_u^{u+\lambda y^2} e^{-(1/2(1-\rho^2))(x-\rho y)^2}\, dx$$
$$\le \frac{\lambda}{2\pi\sqrt{1 - \rho^2}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2}\, dy = \frac{\lambda}{\sqrt{2\pi(1 - \rho^2)}},$$

$$|G_1(u) - \Phi(u)| \le \frac{A}{\sqrt{n}} \int_{-\infty}^{\infty} \left| \frac{x^2 - 1}{\sqrt{\alpha_4 - 1}} \right|^3 dP \le \frac{A\alpha_6}{(\alpha_4 - 1)^{3/2}\sqrt{n}} \text{ by Theorem 2,}$$

$$0 \le K(u) - K_\epsilon(u) \le \epsilon\epsilon(Y^6) \le A\alpha_6\epsilon \text{ by Lemma 12,}$$

$$0 \le L(u) - L_\epsilon(u) \le A\epsilon.$$

Hence

$$\int_{-\infty}^{\infty} \frac{1-\cos Tu}{u^2}\{G(u+\lambda)-\Phi(u+\lambda)\}\,du$$

$$(109) \qquad = \Theta T\left\{\alpha_6\epsilon + \frac{\alpha_6}{(\alpha_4-1)^{3/2}\sqrt{n}} + \frac{1}{\sqrt{n}\sqrt{(\alpha_4-1)(1-\rho^2)}}\right.$$

$$+ \Theta T \iint_{|t_1|\le T} |g(t_1, t_2)|\cdot|f(t_1, t_2) - \varphi(t_1, t_2)|\,dt_1dt_2 .$$

It is easy to verify that

$$\frac{\alpha_6}{(\alpha_4-1)^{3/2}} + \frac{1}{\sqrt{(\alpha_4-1)(1-\rho^2)}} \le \left(\frac{\alpha_6}{\alpha_4 - 1 - \alpha_3^2}\right)^{3/2}.$$

For the left-hand side of (109) we refer to (50) and take x to be the number a therein. Hence

$$T\delta\left\{3\int_0^{T\alpha}\frac{1-\cos u}{u^2}\,du - \pi\right\} \le AT\left\{\alpha_6\epsilon + \frac{1}{\sqrt{n}}\left(\frac{\alpha_6}{\alpha_4-1-\alpha_3^2}\right)^{3/2}\right\}$$

$$(110) \qquad + AT \iint_{|t_1|\le T, |t_2|\le T} |g(t_1, t_2)|\cdot|f(t_1, t_2) - \varphi(t_1, t_2)|\,dt_1dt_2$$

$$+ AT \iint_{|t_1|\le T, |t_2|> T} |g(t_1, t_2)|\,dt_1dt_2 .$$

By Lemma 9 (ii) we have

$$T \iint_{|t_1|\le T, |t_2|> T} |g(t_1, t_2)|\,dt_1\,dt_2$$

$$(111) \qquad \le AT \iint_{|t_1|\le T, |t_2|> T} \frac{1}{|t_2|^3}\left(\lambda + \frac{\lambda^2|t_1|}{\epsilon^{\frac{1}{3}}} + \frac{\lambda^3|t_1|^2}{\epsilon^{\frac{2}{3}}}\right)dt_1\,dt_2 .$$

$$\le A\left(\lambda + \frac{\lambda^2 T}{\epsilon^{\frac{1}{3}}} + \frac{\lambda^3 T^2}{\epsilon^{\frac{2}{3}}}\right).$$

Hence

$$T\delta\left\{3\int_0^{T\delta}\frac{1-\cos u}{u^2}\,du - \pi\right\}$$

$$(112) \qquad \le A\left\{\alpha_6 T\epsilon + \left(\frac{\alpha_6}{\alpha_4-1-\alpha_3^2}\right)^{3/2}\frac{T}{\sqrt{n}} + \lambda + \frac{\lambda^2 T}{\epsilon^{\frac{1}{3}}} + \frac{\lambda^3 T^2}{\epsilon^{\frac{2}{3}}}\right\}$$

$$+ AT \iint_{|t_1|\le T, |t_2|\le T} |g(t_1, t_2)||f - \varphi|\,dt_1\,dt_2 .$$

By Lemma 9 (i) with $k = 3$ we have

$$(113)\qquad T \iint\limits_{|t_1|\le T,|t_2|\le T} |g|\cdot|f-\varphi|\,dt_1\,dt_2 \le \frac{AT\lambda}{\epsilon^{\frac12}} \iint\limits_{|t_1|\le T,|t_2|\le T} |f-\varphi|\,dt_1\,dt_2 .$$

By (37) under Lemma 3,

$$(114)\qquad |f-\varphi| \le \frac{A}{\sqrt{n}}(\beta_{31}|t_1|^3+\beta_{32}|t_2|^3)e^{-\frac14(1-\rho^2)(t_1^2+t_2^2)} \quad \text{for } |t_i| \le \frac{A(1-\rho^2)\sqrt{n}}{\beta_{3i}}$$

with

$$(115)\qquad \begin{aligned} \beta_{31} &= \int_{-\infty}^{\infty}\left|\frac{x^2-1}{\sqrt{\alpha_4-1}}\right|^3 dP \le \frac{4}{(\alpha_4-1)^{\frac32}}\int_{-\infty}^{\infty}(x^6+1)\,dP \\ &\le \frac{8\alpha_6}{(\alpha_4-1)^{\frac32}}, \qquad \beta_{32} = \int_{-\infty}^{\infty}|x|^3\,dP = \beta_3 . \end{aligned}$$

We now take

$$(116)\qquad T = \frac{A}{8}\left(\frac{\alpha_4-1-\alpha_3^2}{\alpha_6}\right)^{\frac32}\sqrt{n},$$

the A coinciding with that in (114). Then

$$(117)\qquad \begin{aligned} \frac{A(1-\rho^2)\sqrt{n}}{\beta_{31}} &\ge \frac{A(1-\rho^2)(\alpha_4-1)^{\frac32}\sqrt{n}}{8\alpha_6} \\ &= \frac{A(\alpha_4-1-\alpha_3^2)\sqrt{\alpha_4-1}\,\sqrt{n}}{8\alpha_6} \ge \frac{A(\alpha_4-1-\alpha_3^2)^{\frac32}\sqrt{n}}{8\alpha_6^{3/2}} = T \end{aligned}$$

$$(118)\qquad \begin{aligned} \frac{A(1-\rho^2)\sqrt{n}}{\beta_{32}} &= \frac{A(\alpha_4-1-\alpha_3^2)\sqrt{n}}{(\alpha_4-1)\beta_3} \\ &\ge \frac{A(\alpha_4-1-\alpha_3^2)^{\frac32}\sqrt{n}}{\alpha_4^{3/2}\beta_3} \ge \frac{A(\alpha_4-1-\alpha_3^2)^{\frac32}\sqrt{n}}{\alpha_6^{3/2}} > T. \end{aligned}$$

Hence (114) is true for $|t_1| \le T$ and $|t_2| \le T$. Using this fact on (113) we obtain

$$(119)\qquad \begin{aligned} T &\iint\limits_{|t_1|\le T,|t_2|\le T} |g||f-\varphi|\,dt_1\,dt_2 \\ &\le \frac{AT\lambda}{\epsilon^{\frac12}}\frac{1}{\sqrt{n}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\frac{\alpha_6}{(\alpha_4-1)^{\frac32}}|t_1|^3+\beta_3|t_2|^3\right\}e^{-\frac14(1-\rho^2)(t_1^2+t_2^2)}\,dt_1\,dt_2 \\ &\le \frac{AT\lambda}{\epsilon^{\frac12}\sqrt{n}}\left(\frac{\alpha_6}{(\alpha_4-1)^{3/2}}+\beta_3\right)\frac{1}{(1-\rho^2)^{5/2}} \\ &= \frac{AT\lambda}{\sqrt{n\epsilon}}(\alpha_6(\alpha_4-1)+\beta_3(\alpha_4-1)^{5/2})\frac{1}{(\alpha_4-1-\alpha_3^2)^{5/2}} \\ &= \frac{AT}{n\sqrt{\epsilon}}(\alpha_6\sqrt{\alpha_4-1}+\beta_3(\alpha_4-1)^2)\frac{1}{(\alpha_4-1-\alpha_3^2)^{5/2}} \\ &\le \frac{AT\alpha_6^{11/6}}{n\sqrt{\epsilon}(\alpha_4-1-\alpha_3^2)^{5/2}} . \end{aligned}$$

Substituting in (112), setting $\epsilon = (\alpha_6 T)^{-1}$ and using (116) we obtain after some easy reduction

$$
\begin{aligned}
&T\delta\left\{3\int_0^{T\delta}\frac{1-\cos u}{u^2}\,du-\pi\right\} \\
(120)\qquad &\le A\left[1+\frac{1}{\sqrt{n(\alpha_4-1)}}+\left(\frac{\alpha_6}{n(\alpha_4-1-\alpha_3^2)}\right)^{\frac{1}{2}}+\left(\frac{\alpha_6}{n(\alpha_4-1-\alpha_3^2)}\right)^{\frac{1}{4}}\right].
\end{aligned}
$$

If $n \ge (\alpha_4 - 1 - \alpha_3^2)^{-1}\alpha_6$, then the right-hand side of (120) is $\le A$, and so, arguing with (120), as we did with (70), we obtain

$$
(121)\qquad \text{l.u.b.}\,|G(u)-\Phi(u)| \le \frac{A}{T} = \frac{A}{\sqrt{n}}\left(\frac{\alpha_6}{\alpha_4-1-\alpha_3^2}\right)^{\frac{1}{2}}.
$$

For $n < (\alpha_4 - 1 - \alpha_3^2)^{-1}\alpha_6$, however, the right-hand side of (121) $\ge A(\alpha_4 - 1 - \alpha_3^2)^{-1}\alpha_6 \ge A$ and (121) becomes a triviality. Hence Theorem 3 is proved.

3.3. To prove Theorem 4, we start again with the identity (103). We have

$$
\begin{aligned}
(122)\qquad K_\epsilon(u) - L_\epsilon(u) &= \{G(u)-H(u)\} - \{G_1(u)-H_1(u)\} \\
&\quad - \{K(u)-K_\epsilon(u)\} + \{L(u)-L_\epsilon(u)\},
\end{aligned}
$$

$$
(123)\qquad 0 \le K(u) - K_\epsilon(u) \le \epsilon\epsilon(Y^{2k}) \le Q_k\epsilon \quad \text{by Lemma 12,}
$$

$$
(124)\qquad 0 \le L(u) - L_\epsilon(u) \le \epsilon\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^{2k}(w(x,y)+|\gamma(x,y)|)\,dx\,dy \le Q_k\epsilon.
$$

Let us show that

$$
(125)\qquad |G_1(u) - H_1(u)| \le Q_k/n^{\frac{1}{2}(k-1)}.
$$

The function $X = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{\xi_i^2-1}{\sqrt{\alpha_4-1}}\right)$ has the same structure as $\sqrt{n}\,\bar{\xi}$ (with $(\alpha_4-1)^{-\frac{1}{2}}(\xi_i^2-1)$ playing the role of ξ_i); hence, by Theorem 1, there exists an asymptotic expansion of the distribution function $G_1(u)$. We shall see that the terms of this asymptotic expansion are precisely $H_1(u)$, whence (125) follows from Theorem 1.

It is obvious that for the polynomial $\psi(it_1, it_2)$ in (35) $\psi(it, 0)$ coincides with the polynomial $\psi(it)$ in (34). Hence the terms of the asymptotic expansion of $G_1(u)$ are the inversion of $e^{-\frac{1}{2}t^2}\{1+\psi(it,0)\}$ viz.

$$
(126)\qquad \Phi(u) + \frac{1}{2\pi}\int_{-\infty}^{u}dx\int_{-\infty}^{\infty}e^{-itx-\frac{1}{2}t^2}\psi(it,0)\,dt.
$$

On the other hand, by (104),

$$
(127)\qquad H_1(u) = \Phi(u) + \int_{-\infty}^{u}dx\int_{-\infty}^{\infty}\gamma(x,y)\,dy,
$$

and by (101) with $t_2 = 0$,

$$
(128)\qquad \int_{-\infty}^{\infty}e^{itx}\,dx\int_{-\infty}^{\infty}\gamma(x,y)\,dy = e^{-\frac{1}{2}t^2}\psi(it,0).
$$

Inversion of (118) gives

$$\int_{-\infty}^{\infty} \gamma(x, y)\, dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx - \frac{1}{2}t^2} \psi(it, 0)\, dt \tag{129}$$

which establishes the equality of $H_1(u)$ and (126).

Using (122), (123), (124), (125) on (103) we get

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1 - \cos Tu}{u^2} \{G(u + z) - H(u + z)\}\, du &= \Lambda_k T \left(\epsilon + \frac{1}{n^{\frac{1}{2}(k-2)}}\right) \\ &+ \Theta T \int\int |g(t_1, t_2)| \cdot |f(t_1, t_2) - \varphi(t_1, t_2)[1 + \psi(it_1, it_2)]|\, dt_1\, dt_2 . \end{aligned} \tag{130}$$

If we expand

$$H(u) = \int\int_{-\lambda y^2 \leq u} \{w(x, y) + \gamma(x, y)\}\, dx\, dy \tag{131}$$

in powers of $n^{-\frac{1}{2}}$ up to and including the term $n^{-\frac{1}{2}(k-3)}$, the remainder is obviously $\Lambda_k n^{-\frac{1}{2}(k-2)}$. Hence

$$H(u) = \Phi(u) + \chi(u) + \Lambda_k / n^{\frac{1}{2}(k-2)}, \tag{132}$$

where $\Phi(u) + \chi(u)$ is the group of terms of the Taylor expansion of (131) in powers of $n^{-\frac{1}{2}}$ up to and including the term $n^{-\frac{1}{2}(k-3)}$. From (130) and (132) we get

$$\begin{aligned} &\left| \int_{-\infty}^{\infty} \frac{1 - \cos Tu}{u^2} \{G(u + z) - \Phi(u + z) - \chi(u + z)\}\, du \right| \\ &\qquad \leq Q_k T \left(\epsilon + \frac{1}{n^{\frac{1}{2}(k-2)}}\right) + AI, \end{aligned} \tag{133}$$

where

$$I = T \int\int_{|t_1| \leq T} |g(t_1, t_2)| \cdot |f(t_1, t_2) - \varphi(t_1, t_2)\{1 + \psi(it_1, it_2)\}|\, dt_1\, dt_2 . \tag{134}$$

We are going to prove that the function $\chi(u)$ here defined satisfies all the requirements of the function $\chi(u)$ in Theorem 4. The structure of $\chi(u)$ announced in Theorem 4 is easily verifiable. It remains to prove the inequalities (15) and (16) satisfied by

$$| G(u) - \Phi(u) - \chi(u) | .$$

It is obvious that the function $\Phi(u) + \chi(u)$ has all the properties of the function $F_1(u)$ in Lemma 8, having a bounded derivative $| \Phi'(u) + \chi'(u) | \leq Q_k$. Hence, on taking z in (133) to be the number a in (51), the left-hand side of (133) does not exceed

$$Q_k T\delta \left(3 \int_0^{T\delta} \frac{1 - \cos u}{u^2}\, du - \pi\right), \qquad \delta = Q_k \,\mathrm{l.u.b.}\, | G(u) - \Phi(u) - \chi(u) |.$$

Hence

$$T\delta\left(3\int_0^{T\delta}\frac{1-\cos u}{u^2}\,du - \pi\right) \le Q_k T\left(\epsilon + \frac{1}{n^{\frac{1}{2}(k-2)}}\right) + Q_k I. \tag{135}$$

In order to appraise I we recall (35) under Lemma 3 (replacing therein each β_{ki} by the larger number $\beta_{k1}\beta_{k2}$, and merging the latter into Q_k)

$$\begin{aligned} |f(t_1, t_2) - \varphi(t_1, t_2)\{1 + \psi(it_1, it_2)\}| \le \frac{Q_k}{n^{\frac{1}{2}(k-2)}}\{\Sigma(|t_i|^k + \cdots \\ + |t_i|^{3(k-2)})\}e^{-(1-\rho^2)(t_1^2+t_2^2)/8} \end{aligned} \tag{136}$$

for

$$|t_i| \le Q_k\sqrt{n}. \tag{137}$$

Put $T = (Q_k\sqrt{n})^l$, with Q_k here coinciding with that in (137) and then (136) is valid for $|t_1| \le T^{1/l}$ and $|t_2| \le T^{1/l}$. Write

$$I = T\iint\limits_{|t_1|\le T^{1/l},\,|t_2|\le T^{1/l}} + T\iint\limits_{|t_1|\le T,\,|t_2|>T^{1/l}} + T\iint\limits_{\substack{T^{1/l}<|t_1|\le T\\ |t_2|\le T^{1/l}}} = I_1 + I_2 + I_3.$$

By Lemma 9 (i),

$$I_1 \le \frac{Q_k T}{n^{\frac{1}{2}}\epsilon^{3/2k}}\iint|f - \varphi(1+\psi)|\,dt_1\,dt_2, \tag{138}$$

whence, by (136)

$$\begin{aligned} I_1 \le \frac{Q_k T}{n^{\frac{1}{2}(k-1)}\epsilon^{3/2k}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sum_{i=1}^{2}(|t_i|^k + \cdots + |t_i|^{3(k-2)})\right) \\ \cdot e^{-(1-\rho^2)(t_1^2+t_2^2)/8}\,dt_1\,dt_2 \le \frac{Q_k T}{n^{\frac{1}{2}(k-1)}\epsilon^{3/2k}}. \end{aligned} \tag{139}$$

By Lemma 9 (iii) we have

$$\begin{aligned} I_2 \le Q_k T\iint\limits_{|t_1|\le T,\,|t_2|>T^{1/l}}\frac{1}{|t_2|^2}\left(\frac{1}{\sqrt{n}\epsilon^{1/2k}}\right. \\ \left.+\frac{|t_1|}{n\epsilon^{3/2k}}\right)\{|f(t_1, t_2)| + \varphi(t_1, t_2)|1 + \psi(it_1, it_2)|\}\,dt_1\,dt_2. \end{aligned}$$

Obviously,

$$\operatorname*{l.u.b.}_{t_2>T^{1/k-2}}\varphi(t_1, t_2)|1 + \psi(it_1, it_2)| = e^{-nQ_k}. \tag{140}$$

On the assumption of non-singularity of $P(x)$ we have, by Lemma 7,

$$\begin{aligned} \operatorname*{l.u.b.}_{|t_2|>T^{1/k-2}}|f(t_1, t_2)| = \operatorname*{l.u.b.}_{|t_2|>Q_k\sqrt{n}}\left|p\left(\frac{t_1}{\sqrt{n}}, \frac{t_2}{\sqrt{n}}\right)\right|^n \\ = \operatorname*{l.u.b.}_{|t_2|\ge Q_k}\left|p\left(\frac{t_1}{\sqrt{n}}, t_2\right)\right|^n = e^{-nQ_k}. \end{aligned} \tag{141}$$

Hence

$$I_2 \leq Q_k T e^{-nQ_k} \iint\limits_{|t_1| \leq T, |t_2| > T^{1/l}} \frac{1}{|t_2|^2} \left(\frac{1}{\sqrt{n}\epsilon^{1/2k}} + \frac{|t_1|}{n\epsilon^{3/2k}} \right) dt_1 \, dt_2 \tag{142}$$

$$= Q_k \left(\frac{n^{l-1}}{\epsilon^{1/2k}} + \frac{n^{(3/2)(l-1)}}{\epsilon^{3/2k}} \right) e^{-nQ_k}.$$

For I_3 we have $|t_1| > T^{1/l} = Q_k\sqrt{n}$, and so Lemma 7 is applicable to I_3 in the same manner as to I_2. Using Lemma 9 (i) on the factor $|g(t_1, t_2)|$ we get

$$I_3 \leq \frac{Q_k n^l e^{-nQ_k}}{\epsilon^{3/2k}}. \tag{143}$$

Combining (135), (138), (139), (142), (143) we obtain

$$T\delta \left(3 \int_0^{T\delta} \frac{1 - \cos u}{u^2} du - \pi \right) \leq Q_k \left(n^{l/2}\epsilon + \frac{n^{l/2}}{n^{\frac{1}{2}(k-2)}} + \frac{n^{l/2}}{n^{\frac{1}{2}(k-1)}\epsilon^{3/2k}} \right) \tag{144}$$

$$+ Q_k \left(\frac{n^{l-1}}{\epsilon^{1/2k}} + \frac{n^{3/2(l-1)}}{\epsilon^{3/2k}} + \frac{n^l}{\epsilon^{3/2k}} \right) e^{-nQ_k}.$$

Putting $\epsilon = \dfrac{1}{n^{k(k-1)/(2k+3)}}$ we get, as the last term in (144) is $\leq Q_k$,

$$T\delta \left(3 \int_0^{T\delta} \frac{1 - \cos u}{u^2} du - \pi \right) \leq Q_k + Q_k n^{l/2} \left(\frac{1}{n^{k(k-1)/(2k+3)}} + \frac{1}{n^{\frac{1}{2}(k-2)}} \right).$$

If $4 \leq k \leq 6$, we take $l = k - 2$ and get

$$T\delta \left(3 \int_0^{T\delta} \frac{1 - \cos u}{u^2} du - \pi \right) \leq Q_k + Q_k \left(\frac{1}{n^{(6-k)/(2(2k+3))}} + 1 \right) \leq Q_k.$$

Hence, by the argument following (70),

$$\text{l.u.b.} \left| G(u) - \Phi(u) - \chi(u) \right| \leq \frac{Q_k}{T} = \frac{Q_k}{n^{\frac{1}{2}(k-2)}},$$

giving (15). If $k \geq 7$, we take $l = \dfrac{2k(k-1)}{2k+3}$ and get

$$T\delta \left(3 \int_0^{T\delta} \frac{1 - \cos u}{u^2} du - \pi \right) \leq Q_k + Q_k \left(1 + \frac{1}{n^{(k-6)/(2(k+3))}} \right) \leq Q_k.$$

Hence

$$\text{l.u.b.} \left| G(u) - \Phi(u) - \chi(u) \right| \leq \frac{Q_k}{T} = \frac{Q_k}{n^{k(k-1)/2(k+3)}},$$

giving (16). Therefore Theorem 4 is proved.

5. When $\alpha_4 - 1 - \alpha_3^2 = 0$. If $\alpha_4 - 1 - \alpha_3^2 = 0$, then there is unit probability that ξ_i assumes exactly two values:

$$Pr\{\xi_i = a\} = p, \qquad Pr\{\xi_i = b\} = q, \qquad p + q = 1.$$

Let $\zeta_i = 1$ with probability p and $\zeta_i = 0$ with probability q. Then $\xi_i = b + (a - b)\zeta_i$, $\eta = (a - b)^2 \frac{1}{n} \Sigma(\zeta_i - \bar{\zeta})^2$. Hence it is sufficient to consider the variable $\frac{1}{n} \sum (\zeta_i - \bar{\zeta})^2 = \eta$. Letting $\Sigma\zeta_i = r = np + \sqrt{npq}\, X$ we have $\eta_1 = r - \frac{r^2}{n} = npq + (q - p)\sqrt{npq}\, X - pqX^2$. We now consider two distinct cases:

Case (*i*). $p \neq q$. Here

$$F(z) = Pr\left\{\frac{\eta_1 - n/\partial q}{|p - q|\sqrt{npq}} \leq z\right\}$$
$$= Pr\{(X + c\sqrt{n})^2 \geq c^2 n - 2|c|\sqrt{n}\, z\}, \quad c = \frac{p - q}{2\sqrt{pq}}.$$

Thus $F(z) = 1$ if $z \geq \frac{1}{2} | c | \sqrt{n}$. If $z < \frac{1}{2} | c | \sqrt{n}$, then

$$F(z) = Pr\{X \leq -cn - (c^2 n - 2 | c | \sqrt{n} z)^{\frac{1}{2}}\}$$
$$+ Pr\{X \geq -c\sqrt{n} + (c^2 n - 2 | c | \sqrt{n} z)^{\frac{1}{2}}\} = F_1(z) + F_2(z).$$

To the random variable X Theorem 2 can be applied. Suppose that $c < 0$; then, by Tchebycheff's inequality,

$$F_2(z) \leq Pr\{X \geq -cn\} \leq \frac{1}{c^2 n} \leq \frac{1}{(p - q)^2 n}.$$

By Theorem 2,

$$F_1(z) = Pr\{X \leq -cn - (c^2 n - 2|c|\sqrt{n}\, z)^{\frac{1}{2}}\}$$
$$= \Phi(z) + \frac{\Theta z^2}{\sqrt{n}\,|p - q|} + \frac{\Theta(p^2 + q^2)}{\sqrt{npq}}.$$

Hence

$$(145) \qquad |F(z) - \Phi(z)| \leq A\left\{\frac{p^2 + q^2}{\sqrt{npq}} + \frac{z^2}{\sqrt{n}\,|p - q|} + \frac{1}{n(p - q)^2}\right\}.$$

The same inequality holds also for $c > 0$.

Case (*ii*). $p = q = 1/2$. Here $\eta_1 = \frac{1}{4}(n - X^2)$; hence

$$(146) \qquad Pr\left\{\eta_1 \geq \frac{n - z}{4}\right\} = Pr\{X^2 \leq z\} = \frac{1}{\sqrt{2\pi}} \int_0^z x^{-\frac{1}{2}} e^{-x/2}\, dx + \frac{\Theta}{\sqrt{n}}.$$

There is no asymptotic expansion for the distribution function of η_1. (See (C), p. 83.)

SAMPLING INSPECTION PLANS FOR CONTINUOUS PRODUCTION WHICH INSURE A PRESCRIBED LIMIT ON THE OUTGOING QUALITY

A. WALD AND J. WOLFOWITZ

Columbia University

1. Introduction. This paper discusses several plans for sampling inspection of manufactured articles which are produced by a continuous production process, the plans being designed to insure that the long-run proportion of defectives shall not exceed a prescribed limit. The plans are applicable to articles which can be classified as "defective" or "non-defective" and which are submitted for inspection either continuously or in lots. In Section 2 the notions of "average outgoing quality limit" and "local stability" are discussed. The valuable concept of average outgoing quality limit for lot inspection is due to Dodge and Romig [4], and that for inspection of continuous production to Dodge [1]. Section 3 contains a description of a simple inspection plan (SPA) applicable to to continuous production and a proof that the plan will insure a prescribed average outgoing quality limit. Section 4 contains a proof that this inspection plan also has the important property that it requires minimum inspection when the production process is in statistical control. In Section 5 is contained the description of a general class of plans which possess both these important properties.

The problem of adapting SPA to the case when the articles are submitted for inspection in lots instead of continuously, is treated in Section 6. Some methods of achieving local stability are discussed in Section 7 and a specific plan is developed there. Finally Section 8 discusses the relationship between the present work and that of the earlier and very interesting paper of H. F. Dodge [1], mentioned above.

If a quick first reading is desired the reader may omit the second half of Section 3 (which contains a proof of the fact that SPA guarantees the prescribed average outgoing quality limit) and the entire Section 4 except for its title (the proof of the statement made in the title of Section 4 occupies the whole section).

2. Fundamental notions. In this paper we shall deal only with a product whose units can be classified as "defective" or "non-defective." We shall assume that the units of the product are submitted for inspection continuously, except in Section 6, where we assume that they are submitted in lots. Throughout the paper we shall assume that the inspection process is non-destructive, that it invariably classifies correctly the units examined, and that defective units, when found, are replaced by non-defectives. By the "quality" of a sequence of units is meant the proportion of defectives in the sequence as produced. By the "outgoing quality" (OQ) of a sequence is meant the proportion of defectives after whatever inspection scheme which is in use has been applied. If this scheme involves random sampling, then in general the OQ is a chance variable.

(It depends on the variations of random sampling.) If the OQ converges to a constant p_a with probability one as the number of units produced increases indefinitely, p_a is called the "average outgoing quality" (AOQ). The AOQ when it exists is therefore the average quality, in the long run, of the production process after inspection. It is a function of both the production process and the inspection scheme. These definitions are due to Dodge [1].

The "average outgoing quality limit" (AOQL) is a number which is to depend only on the inspection scheme and not at all on the production process. Roughly speaking, it is a number, characteristic of an inspection scheme, such that no matter what the variations or eccentricities of the production process, the AOQ never exceeds it. For the purposes of this paper we shall need the following precise definition: Let c_i be zero or one according as the ith unit of the product, before application of the inspection scheme, is a non-defective or a defective, respectively. Let d_i have a similar definition *after* application of the inspection scheme. (We note that if the ith item was inspected, then $d_i = 0$; if the ith item was not inspected, then $c_i = d_i$.) The sequence $c = c_1, c_2, \cdots, c_N, \cdots$, ad inf. characterizes the production process[1]. The elements of $d = d_1, d_2, \cdots$, ad inf. are in general chance variables. The number L is called the AOQL if it is the smallest[2] number with the property that the probability is zero that

$$\limsup_N \frac{\sum_{i=1}^N d_i}{N} > L,$$

no matter what the sequence c.

It should be noted that this definition of AOQL places no restrictions whatever on the production process, since *all* sequences c are admitted. It is too much to expect a production process to remain always in control; indeed, doubt as to whether statistical control always exists may cause a manufacturer to institute an inspection scheme. The inspection schemes which we shall give below will yield a specified AOQL no matter what the variations in production are. If these schemes are employed, then, even if Maxwell's demon of gas theory fame were to transfer his activities to the production process, he would be unsuccessful in an effort to cause the AOQL to be exceeded. A dishonest manufacturer might sometimes essay to do this. If we imposed restrictions on the sequence c and

[1] This use of an infinite sequence to describe the production process deserves a few words. What we consider in this paper are schemes applicable when the number of units produced is large and operate mathematically as if the production sequence were of infinite length. Naturally the latter is never the case in actuality. However, the larger the number of units produced the more nearly will the reality conform to the results derived from the mathematical model. While the present definition uses explicitly the notion of an infinite sequence, such a commonplace statement as "the probability is 1/2 that a coin will fall heads up" uses this notion implicitly. It is also implicit in the intuitive meaning we ascribe to such a word as "average," which is in every day use.

[2] It is not difficult to see that such a number always exists, for it is the lower bound of a set which is non-empty (it contains the point one), bounded from below (zero is a lower bound), and closed.

determined the AOQL on that basis, we would run the danger that the relative frequency of defects in the sequence of outgoing units might exceed the AOQL if it happened that the actual sequence c did not satisfy the restrictions imposed.

After we discuss below various possible sampling inspection plans which insure that the AOQL does not exceed a predetermined value L, it will be seen that for any given $L > 0$ there are many sampling inspection schemes which do this. To choose a particular sampling plan from among them the following considerations may be advanced: If two inspection plans S and S' both insure the inequality AOQL $\leq L$ and if for any sequence c the average number of inspections required by S is not greater than that required by S' and if for some sequences c the average number of inspections required by S is actually smaller than that required by S', then S may be considered, in general, a better inspection plan than S'. However, the amount of inspection required by a sampling plan is not always th *only* criterion for the selection of a proper sampling scheme. There may be also other features of a sampling plan which make it more or less desirable. We shall mention here one such feature, called "local stability," which will play a role in our discussions later. Consider the sequence d obtained from the sequence c by applying a sampling inspection scheme. Even if the AOQL does not exceed L, it may still happen that there will be many large segments of the sequence d within which the relative frequency of ones is considerably higher than L. For instance, it may happen that in the segment $(d_1, \cdots, d_m)$ the relative frequency of ones is equal to $\frac{3}{2}L$, in the segment $(d_{m+1}, \cdots, d_{2m})$ the relative frequency is equal to $\frac{1}{2}L$, in the segment $(d_{2m+1}, \cdots, d_{3m})$ the relative frequency is again equal to $\frac{3}{2}L$, and this is followed again by a segment of m elements where the relative frequency of ones is equal to $\frac{1}{2}L$, and so forth. If m is large, such a sequence d is not very desirable, since each second segment will contain too many defects. A sequence d is said to be not locally stable if there exists a large fixed integer m such that the relative frequency of ones in $(d_{k+1}, \cdots, d_{k+m})$ is considerably greater than L for *many* integral values k. On the other hand, the sequence d is said to be locally stable if for any large m the relative frequency of ones in $(d_{k+1}, \cdots, d_{k+m})$ is not substantially above L for nearly all integral values k. This is clearly not a precise definition of "local stability," but merely an intuitive indication of what we want to understand by the term, since we did not define what we mean by "large m," "many values of k," "considerably above L," etc. A precise definition of local stability will not be needed in this paper, since it is not our intention to develop a complete theory for the choice of the sampling plan. The idea of local stability will be used in this paper merely for making it plausible that some schemes we shall consider behave reasonably in this respect. A similar idea, called "protection against spotty quality," is discussed by Dodge [1]. A possible precise definition of local stability could be given in terms of the frequency with which $F(N) = \frac{1}{(k+1)} \sum_{i=N}^{N+k} d_i$ (k being fixed) lies within given limits.

3. A sampling inspection plan which insures a given AOQL no matter what the variations in the production process. The only feature of the sampling (inspection) plan (SP) studied in this section and hereafter referred to as SPA which we shall consider here is that it insures the achievement of a specified AOQL. Considerations leading to a choice among several schemes are postponed to later sections.

For convenience, let f be the reciprocal of a positive integer. SPA calls for alternating partial inspection and complete inspection. Partial inspection is performed by inspecting one element chosen at random from each of successive groups of $\frac{1}{f}$ elements. Complete inspection means the inspection of every element in the order of production. SPA is completely defined when a rule is given for ending one kind of inspection and beginning the other.

It is clear that all SP need not be of the above class. Thus, for example, a scheme might consist of partial inspection with various f's employed in various sequences. We make no attempt in this paper to examine all possible schemes. For simplicity in practical operation, alternation of complete inspection and partial inspection with fixed f would seem reasonable. The Dodge scheme [1] is of this type.

We shall also not discuss the question of a choice of the constant f, but will assume that a particular value has been chosen for various reasons and is a datum of our problem. Reasons which might influence a manufacturer in his choice of f could be contract specifications which impose a minimum on the amount of inspection, or psychological grounds to the same effect. The manufacturer may desire a certain minimum amount of inspection in order to detect malfunctioning of his production process. Also f controls local stability to some extent. The consequences of a choice of f as they appear in the theory below may also play a role.

Returning to SPA, we begin with partial inspection. Let L be the specified AOQL. Denote by k_N the number of groups of $\frac{1}{f}$ units in which defectives were found as the result of *partial* inspection from the beginning of production through the Nth unit. SPA is as follows:

(a) Begin with partial inspection.

(b) Begin full inspection whenever

$$e_N = \frac{k_N\left(\frac{1}{f} - 1\right)}{N} > L.$$

(c) Resume partial inspection when

$$e_N \leq L.$$

(d) Repeat the procedure. (It will be recalled that defective units, when found, are always to be replaced with non-defectives.)

It is to be observed that in this plan the number of partial inspections increases without limit. For, while complete inspection is going on, the value of k_N remains constant, so that after a long enough period of complete inspection the denominator N of the expression which defines e_N will have increased sufficiently for e_N to be not greater than L. On the other hand, complete inspection may never occur. This will be the case if, for example, no defectives or very few defectives are produced.

We shall now show that the AOQL of the above SP is L. We first note that, at N, e_N can increase only by $\dfrac{\left(\frac{1}{f} - 1\right)}{N}$. Hence, for sufficiently large N, $e_N < L + \epsilon$, where $\epsilon > 0$ may be arbitrarily small.

Suppose now that the production process is subject to any variations whatsoever, i.e., the sequence

$$c = c_1, c_2, \cdots, c_N, \cdots, \text{ ad inf.}$$

is any arbitrary sequence whatever (by their definition the c_i are all zero or one). Our result is therefore proved if we show that, with probability one,

$$\lim_{N\to\infty} \left(e_N - \frac{1}{N}\sum_{i=1}^{N} d_i\right) = 0 \tag{3.1}$$

for this arbitrary c, and that for at least one c

$$\lim_{N\to\infty} e_N = L. \tag{3.2}$$

Let $S(N)$ be the number of groups of $\frac{1}{f}$ units which have been partially inspected through the Nth unit. Define x_i as zero if in the ith partially inspected group a non-defective was found and as one if a defective was found. We have

$$k_N = \sum_{i=1}^{S(N)} x_i .$$

Since the number of times partial inspection takes place increases indefinitely, $S(N) \to \infty$ as $N \to \infty$. Also $S(N) \leq fN < N$. Let α_j be the serial number of the last unit in the jth partially inspected group. Then for all j the expected value $E(x_j)$ of x_j is given by

$$E(x_j) = f\left(\sum_{i=(\alpha_j-(1/f)+1)}^{\alpha_j} c_i\right).$$

We have, for all j

$$\sum_{i=(\alpha_j-(1/f)+1)}^{\alpha_j} (c_i - d_i) = x_j \tag{3.3}$$

so that

$$E\left(\left[\frac{1}{f} - 1\right] x_j - \sum_{\alpha_j-(1/f)+1}^{\alpha_j} d_i\right) = 0.$$

Also from (3.3) it follows, since x_j is the value of a binomial chance variable from a population of fixed number $\left(\frac{1}{f}\right)$, that there exists a positive constant β such that

$$\sigma^2\left(\left[\frac{1}{f} - 1\right]x_j - \sum_{\alpha_j-(1/f)+1}^{\alpha_j} d_i\right) < \beta \tag{3.4}$$

where $\sigma^2(x)$ is the variance of a chance variable x. Now a theorem of Kolmogoroff (Kolmogoroff [2], Fréchet [3], p. 254) states:
A sequence of chance variables with zero means and variances $\sigma_1^2, \sigma_2^2, \cdots$ converges with probability one towards zero in the sense of Cèsaro if

$$\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} \tag{3.5}$$

converges. The inequality (3.4) permits us to apply this theorem to the sequence of chance variables of which the jth ($j = 1, 2, \cdots$ ad inf.) is

$$\left(\left[\frac{1}{f} - 1\right]x_j - \sum_{\alpha_j-(1/f)+1}^{\alpha_j} d_i\right),$$

since the series $\sum_{i=1}^{\infty} \frac{1}{i^2}$ is well known to be convergent. We therefore obtain that, with probability one,

$$\lim_{S(N)\to\infty} \frac{\left(\left[\frac{1}{f} - 1\right]\sum_{j=1}^{S(N)} x_j - \sum_{j=1}^{N} d_j\right)}{S(N)} = \lim_{N\to\infty} \frac{N}{S(N)}\left(e_N - \frac{1}{N}\sum_{i=1}^{N} d_i\right) = 0,$$

since the units which are fully inspected contribute nothing to Σd_i. Since $S(N) < N$, the desired result (3.1) is a fortiori true.

If c is such that all the c_i are one, it is readily seen that (3.2) holds. If many (this adjective can be precisely defined) defectives are produced, this will also be the case. This completes the proof of the fact that the AOQL of SPA is L no matter how capriciously the production process may vary.

4. When the production process is in statistical control, SPA requires minimum inspection. The production process is said to be in statistical control if there is a positive constant $p \leq 1$ such that, for every i, the probability that $c_i = 1$ is p and is independent of the values taken by the other c's. We shall see that if the process is in statistical control and if SPA is applied to it, the specified AOQL is guaranteed with a minimum amount of inspection.

The number of units inspected through the Nth unit produced is

$$I(N) = N - \left(\frac{1}{f} - 1\right)S(N). \tag{4.1}$$

If the process is in statistical control we have, with probability one,

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} c_i}{N} = p \tag{4.2}$$

by the strong law of large numbers. Shortly we shall prove the existence of a constant L^* such that, with probability one,

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} d_i}{N} = L^*. \tag{4.3}$$

Assume for the moment that this is so. Since it is only by inspection that defectives are removed, and the units selected for inspection are in statistical control like the original sequence, it follows that, with probability one,

$$\lim_{N\to\infty} \frac{I(N)}{N} = \frac{1}{p}(p - L^*) = 1 - \frac{L^*}{p} \tag{4.4}$$

because, with probability one,

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} (c_i - d_i)}{N} = p - L^*.$$

Inspection is therefore at a minimum when L^* is at a maximum compatible with the specified AOQL. By (4.3) the latter means that

$$L^* \leq L. \tag{4.5}$$

SPA has been shown to guarantee this requirement. The optimum situation from the point of view of the amount of inspection would therefore be to have $L^* = L$, but this cannot always be achieved. The absolute minimum amount of inspection clearly is f, i.e., partial inspection exclusively. Consequently from (4.4)

$$1 - \frac{L^*}{p} \geq f$$

so that

$$L^* \leq p(1 - f). \tag{4.6}$$

Combining (4.5) and (4.6) we see that we have to consider three cases:

Case a. If

$$p > \frac{L}{1-f} \tag{4 7}$$

we have to show that

$$L = L^*. \tag{4.8}$$

Case b. If

$$p < \frac{L}{1-f} \tag{4.9}$$

we have to show, by (4.4), that

$$1 - \frac{L^*}{p} = f,$$

that is,

$$L^* = p(1 - f). \tag{4.10}$$

Case c. If

$$p = \frac{L}{1 - f} \tag{4.11}$$

we have to show that

$$L = L^* = p(1 - f). \tag{4.12}$$

PROOF of (4.8): We have already remarked in Section 3 that in SPA partial inspection always recurs, but complete inspection need never occur. We shall show in a moment that (4.7) implies that no matter how large an integer γ is chosen, the probability of temporarily stopping partial inspection for some $N > \gamma$ is one. Assume that this is so. Choose an arbitrarily small positive ϵ, and let $\gamma > \dfrac{\left(\frac{1}{f} - 1\right)}{\epsilon}$. For a sequence where complete and partial inspection alternate infinitely many times let

$$A = \alpha_1, \alpha_2, \cdots, \text{ad inf.}$$

be the sequence of integers at which partial inspection ends, and let

$$B = \beta_1, \beta_2, \cdots, \text{ad inf.}$$

be the sequence of integers at which complete inspection ends. Then, for all j,

$$\alpha_{j+1} > \beta_j > \alpha_j .$$

From the description of SPA it follows that, for all $N > \gamma$ which belong to either A or B,

$$| e_N - L | < \epsilon. \tag{4.13}$$

In Section 3 we proved

$$\lim_{N\to\infty} \left(e_N - \frac{1}{N} \sum_{i=1}^{N} d_i \right) = 0 \tag{3.1}$$

with probability one. Since ϵ is arbitrarily small it follows that, with probability one,

$$\lim_{\substack{N\to\infty \\ (N \text{ in } A \text{ or } B)}} \frac{\sum_{i=1}^{N} d_i}{N} = L. \tag{4.14}$$

To complete the proof of (4.8) we have still to show that L^* exists and that the probability is one that complete inspection will occur infinitely many times. First we prove that L^* exists.

As N increases during an interval of complete inspection, $D(N) = \sum_{i=1}^{N} d_i$ remains constant. Hence $\frac{D(N)}{N}$ decreases monotonically. Since for the ends of such intervals (4.14) holds, it follows that (4.14) holds as $N \to \infty$ and is a member of A, B, or an interval (α_j, β_j) for all j.

Let $N \to \infty$ while always being in the interior of an interval $(\beta_j, \alpha_{j+1}]$, $j = 1, 2, \cdots$, ad inf., which contains α_{j+1} but not β_j. Let N^* be the total number of units in these intervals through the Nth unit produced. Let N_1 and N_2 be such that

$$\beta_j = N_1 < N_2 < \alpha_{j+1}.$$

Then

$$N_2^* - N_1^* = N_2 - N_1.$$

Since the production process is in statistical control, we have, by the strong law of large numbers,

$$\lim_{N\to\infty} \frac{D(N)}{N^*} = p(1 - f) = p' \tag{4.15}$$

with probability one. Let δ^* be the general designation for numbers $<\epsilon$ in absolute value, so that all δ^* are not the same. With probability one for almost all N, we have by (4.15)

$$\frac{D(N_1)}{N_1^*} = p' + \delta^*$$

$$\frac{D(N_2)}{N_2^*} = p' + \delta^*.$$

Write

$$\frac{[D(N_2) - D(N_1)]}{(N_2 - N_1)} = K.$$

Now

$$\frac{D(N_2)}{N_2^*} = \frac{D(N_1) + [D(N_2) - D(N_1)]}{N_1^* + (N_2^* - N_1^*)} = \frac{D(N_1) + [D(N_2) - D(N_1)]}{N_1^* + (N_2 - N_1)}$$

$$= \frac{(p' + \delta^*)N_1^* + K(N_2 - N_1)}{N_1^* + (N_2 - N_1)} = p' + \delta^*.$$

Hence

$$K(N_2 - N_1) = 2\delta^* N_1^* + (p' + \delta^*)(N_2 - N_1). \tag{4.16}$$

Now suppose (4.3) does not hold. From the definition of AOQL it follows that for some $\eta > \epsilon$ there exist sequences (whose totality has a positive probability) so that, for infinitely many N_2 we have

$$\frac{D(N_2)}{N_2} = \frac{D(N_1) + [D(N_2) - D(N_1)]}{N_1 + (N_2 - N_1)} < L - 4\eta. \tag{4.17}$$

For large enough N_1, from (4.14),

$$\frac{D(N_1)}{N_1} = L + \delta^*$$

with probability one and hence, using (4.16) in (4.17)

$$\begin{aligned} N_1(L + \delta^*) + 2\delta^* N_1^* + (p' + \delta^*)(N_2 - N_1) \\ < LN_1 + L(N_2 - N_1) - 4\eta N_2 \end{aligned} \tag{4.18}$$

from which, using the fact that $p' \geq L$ (from (4.7)), we get

$$N_1\delta^* + 2N_1^*\delta^* + \delta^*(N_2 - N_1) < -4\eta N_2 . \tag{4.19}$$

((4.18) and (4.19) hold for the sequences for which (4.17) holds, except perhaps on a set of sequences whose probability is zero.) Since $N_1^* \leq N_1$ and $| \delta^* | < \eta$, we have, on the other hand,

$$\begin{aligned} N_1\delta^* + 2N_1^*\delta^* + \delta^*(N_2 - N_1) &\geq -3\eta N_1 - \eta(N_2 - N_1) \\ &> -4\eta N_1 - 4\eta(N_2 - N_1) = -4\eta N_2 \end{aligned} \tag{4.20}$$

which contradicts (4.19) and proves the desired result ((4.3) and (4.8)), except that it remains to prove that, no matter how large γ, the probability of temporarily stopping partial inspection at some $N > \gamma$ is one. Let $\gamma_0 \geq \gamma$ be some integer at which partial inspection is going on. From (4.2) and (4.7) it would follow, if partial inspection never ceased on a set of sequences with positive probability, that, on this set, with conditional probability one, for N sufficiently large and ϵ sufficiently small,

$$\frac{k_N - k_{\gamma_0}}{f(N - \gamma_0)} > \frac{L}{1 - f} + \epsilon,$$

$$\frac{N}{N - \gamma_0} \frac{k_N(1 - f)}{fN} > L + (1 - f)\epsilon,$$

$$e_N > L \frac{N - \gamma_0}{N} + \frac{(N - \gamma_0)(1 - f)\epsilon}{N},$$

$$e_N > L + \frac{(1 - f)\epsilon}{2}.$$

This contradiction proves that complete inspection is eventually resumed and completes the proof of minimum inspection in Case a.

Proof of (4.10): We shall prove that (4.9) implies that, with probability one, complete inspection will cease, never to be resumed. For, from (4.15) and (4.9) it follows that for N sufficiently large and ϵ sufficiently small,

$$\frac{D(N)}{N^*} = p' + \delta^* < L - 2\epsilon. \tag{4.21}$$

Hence, a fortiori,

$$\frac{D(N)}{N} < L - 2\epsilon. \tag{4.22}$$

((4.21) and (4.22) hold with probability one.)
(3.1) states that, with probability one,

$$\lim_{N\to\infty}\left(e_N - \frac{D(N)}{N}\right) = 0.$$

Hence for all N sufficiently large, with probability one,

$$e_N < L - \epsilon,$$

i.e., with probability one complete inspection is never resumed.

When (4.9) holds, therefore, with probability one and with a finite number of exceptions SPA will require only partial inspection.

Proof of (4.12): If $p = \frac{L}{1-f}$ and complete inspection finally never resumes, then (4.12) follows easily. If $p = \frac{L}{1-f}$ and partial and complete inspection alternate infinitely many times, then the proof is similar to that of (4.8) and is therefore omitted. In either case the desired result follows.

5. A class of SP all of which insure both a given AOQL and minimum inspection. Let the definition of SPA be modified in the following particulars:

(b) Begin full inspection whenever

$$e_N = \frac{k_N\left(\frac{1}{f} - 1\right)}{N} > L + \phi(N).$$

(c) Resume partial inspection when

$$e_N \leq L - \psi(N).$$

Let $\phi(N)$ and $\psi(N)$ be such that

$$-\psi(N) \leq \phi(N)$$

$$\lim_{N\to\infty} \phi(N) = \lim_{N\to\infty} \psi(N) = 0.$$

(SPA corresponds to the case $\phi(N) \equiv \psi(N) \equiv 0$.) Then all the SP of this class have the property that the AOQL is L and that inspection is at a minimum in

the sense of Section 4. The proofs are essentially the same as those for SPA and hence will be omitted.

6. The inspection plans of Section 5 can also be applied to lot inspection. We shall carry on the discussion of this section in terms of SPA, but the results apply to all the members of the class of plans described in Section 5. We shall show that SPA can also be applied when the product is submitted for inspection in lots. Although we assumed previously that the units of the product are arranged in order of production, the results obtained for SPA remain valid for any arbitrary arrangement of the units. If the product is submitted in lots we may arrange the units as follows: Let $l_1, l_2, \cdots$, etc. be the successive lots in the order of their submission for inspection. Within each lot we consider the units arranged in the order in which they are chosen for inspection. In this way we have arranged all units in an ordered sequence and the inspection can be applied as described before. Thus, we start with partial inspection, i.e., we take out groups of $\frac{1}{f}$ elements in l_1 and inspect one unit (selected at random) from each of these groups. When $e_N > L$, we start complete inspection and revert to partial inspection as soon as $e_N \leq L$. When the units in l_1 are used up in the process of inspection, we continue, using the units of l_2, etc.

If it is found inconvenient to take out a group of $\frac{1}{f}$ units and then to select one unit for inspection, we could modify the sampling inspection plan as follows: Instead of taking out a group of $\frac{1}{f}$ units and then selecting at random one unit from it, we select at random *one* unit from the uninspected part of the lot and look upon this unit as the unit selected at random from a hypothetical group of $\frac{1}{f}$ units. Thus we can proceed exactly as before, except that we have to keep in mind that with each unit inspected under "partial inspection" we have used up another set of $\frac{1}{f} - 1$ units. Thus, as soon as $\left(\frac{1}{f} - 1\right)$ times the number of units inspected under "partial inspection" becomes equal to or greater than the number of units in the uninspected part of the lot, the inspection of that lot is already terminated, and we have to start using the units of the next lot. The inconvenience caused by the necessity of keeping track of the number of units inspected under "partial inspection" and of the number of units in the uninspected part of the lot can be eliminated by further modifying the inspection plan as follows: Instead of beginning complete inspection as soon as $e_N > L$, we continue "partial inspection" until $E_N = e_N - L$ is so large that complete inspection of all the units of the lot not yet used up has to be made in order to bring e_N down to L at the end of the lot. This leads to the following sampling procedure, to be known as SPB: Let N_0 be the number of units in the lot, let N_L be the serial number of the last unit in the preceding lot, and let $E(N_L) =$

$N_L E_{N_L} = N_L(e_{N_L} - L)$ be the "excess" carried over from the preceding lot. For simplicity assume that the following are all integers:

$$LN_0 = M$$

$$\frac{fM}{1-f} = M^*$$

$$fN_0 = N^*$$

and

$$\frac{fE(N_L)}{1-f} = E^*.$$

The inspection procedure is then as follows: Inspect successive units drawn at random until either

(a) $M^* - E^*$ defectives have been found in the first $N' < N^*$ units inspected. In this case inspect further an additional $N_0 - \dfrac{N'}{f}$ units and this terminates the inspection of the lot. The excess to be carried over to the next lot is then zero.

Or

(b) N^* units have been inspected and the number of defectives found is $H \leq M^* - E^*$. In this case the inspection of the lot is terminated and the present negative excess

$$E(N_L + N_0) = [H - (M^* - E^*)]\,\frac{(1-f)}{f}$$

is carried over to the next lot. (The serial number of the last element in the present lot is $N_L + N_0$ and

$$e_{(N_L+N_0)} = \frac{N_L e_{N_L} + H\,\dfrac{(1-f)}{f}}{N_L + N_0}.$$

Hence the present excess is

$$\begin{aligned}(N_L + N_0)[e_{(N_L+N_0)} - L] &= N_L e_{N_L} + H\,\frac{(1-f)}{f} - LN_L - LN_0\\ &= N_L(e_{N_L} - L) + H\,\frac{(1-f)}{f} - M\\ &= \frac{(1-f)}{f}\,[H - M^* + E^*],\end{aligned}$$

as given above.)

We note an important property of SPB: The excess carried over from a preceding lot is never positive.

7. Possible modifications of the SP to achieve local stability. Although the sampling plans discussed in previous sections are optimum in the sense that they guarantee the desired AOQL with a minimum of inspection when the production process is in statistical control, they do not always behave very favorably as far as local stability is concerned. To make this point clear, consider the following example: Suppose that during a very long initial time period the production process functions very well and the relative frequency of defectives produced is well below L. Thus, applying SPA, say, $e_N - L$ will be considerably less than zero at the end of this period. Now suppose that then the production process suddenly deteriorates and the number of defectives produced during the next period of time is considerably higher than L. In spite of that, complete inspection will not begin for quite some time because e_N became so small during the initial period. Thus there will be a long segment in the sequence of outgoing units within which the relative frequency of defectives will be larger than the prescribed AOQL. Of course, this segment will be counterbalanced by other segments where the relative frequency of defectives will be below the AOQL, so that the AOQL will not be violated. Nevertheless, the occurrence of long segments with too many defectives, i.e., a lack of local stability, is not desirable.

It should be noted that, even though SPA was not designed to achieve considerable local stability, drastic lack of local stability cannot occur when the production process is in statistical control and SPA is employed. In the example given above where the outgoing quality was not locally stable, it was assumed that there were variations in the production process. The existence of statistical control acts as an important stabilizing factor on the quality.

In this section we want to discuss several possible modifications of SPA which will insure a greater degree of local stability. One such modification is the following: We choose a positive constant A and we define the excess E_N^* for each value N as follows: $E^*(N)$ is equal to the excess $E(N)$ as originally defined ($= N[e_N - L]$) as long as for all $N' \leq N$, $E(N') \geq -A$. The difference $E^*(N+1) - E^*(N) = E(N+1) - E(N)$ for all N for which $E(N+1) - E(N) \geq 0$. If $E(N+1) - E(N) < 0$, then $E^*(N+1) = \max\,[E^*(N) + \{E(N+1) - E(N)\}, -A]$. In other words, with this modification of the sampling inspection plan we set a lower bound $-A$ for the excess. When the excess is positive we begin complete inspection, and revert to partial inspection when the excess becomes non-positive. The effect of this is that, if the proportion of defectives produced becomes large, complete inspection will not be delayed very long, although the proportion of defectives produced in the preceding period may have been considerably below L. It is clear that this modification of SPA does not increase the AOQL. However, the amount of inspection will be somewhat increased, especially when the quality of the product is less than or only slightly greater than L. If the constant A is large, the increase in the amount of inspection is only slight, but also the degree of local stability achieved is not very high. On the other hand, if A is small, the increase

in the amount of inspection may be considerable, but a high degree of local stability is achieved. Thus, the choice of A should be made so that a proper balance between local stability and amount of inspection is achieved.

Modifying SPA by setting a lower limit for the excess has the disadvantage that the mathematical treatment of this case is involved. We shall, therefore, consider another modification of the inspection plan which will have largely the same effect, but whose mathematical treatment appears to be much simpler. A fixed positive integer N_0 is chosen and the inspection scheme is designed so that $E_{N_0} \leq 0$ is assured. If E_{N_0} is negative, we replace it by zero. In other words, no excess is carried over from the first segment of N_0 units to the next segment of N_0 units. Thus, the second segment of N_0 units is treated exactly the same way as if it were the first segment, and this is repeated for each consecutive segment of N_0 units. This modification of SPA (the resulting plan is to be known as SPC) has essentially the same effect as setting a lower bound for the excess. Again it is clear that by this modification the AOQL is not increased, but the amount of inspection may be increased. The latter is particularly true when N_0 is small, which corresponds to very high local stability requirements. More efficient plans than SPC can probably be devised for this situation.

Undoubtedly, there are many other possible modifications of the inspection plan by which a greater degree of local stability can be achieved at the price of somewhat increased inspection. It is not the purpose of this paper to enumerate all these possibilities or to develop a theory as to which of them may be considered an optimum procedure. We shall restrict ourselves to a discussion of the mathematical consequences of SPC. First we define it precisely. If it is to be applied to inspection of lots of size N_0 then SPC is simply SPB with $E(N_L)$ and E^* always zero. When applied to continuous production it will operate as follows: Assume for convenience that $M = LN_0$, $N^* = fN_0$, and $\frac{fM}{1-f} = M^*$ are all integers.

(a) Begin each segment of N_0 units with partial inspection, i.e., inspect one unit chosen at random from each successive group of $\frac{1}{f}$ units. Continue partial inspection until one of the following events occurs: either

(b) M^* defectives are found. In this case begin complete inspection with the first unit which follows the group in which the last of the M^* defectives was found and continue until the end of the segment of N_0 units.

or

(b′) N^* groups of $\frac{1}{f}$ units are partially inspected.

(c) Repeat with the next segment of N_0 units.

Comparison with SPB shows that, in SPC, if (b) occurs earlier or at the same time as (b′), then $E_{N_0} = 0$, while if (b′) occurs before (b) we have $E_{N_0} < 0$. In contradistinction to SPB, in SPC there is no carrying over of the excess.

Let us determine the AOQ for SPC when the production process is in a state

of statistical control. Denote by p the probability that a unit produced will be defective. Let the chance variable H denote the number of defectives found during partial inspection. The probability that $H = i < M^*$ is

$$\binom{N^*}{i} p^i (1-p)^{N^*-i}.$$

$H \leq M^*$ always. We have, when $H = i$,

$$E(N_0) = \frac{(1-f)i}{f} - LN_0,$$

and hence

$$N_0 e_{N_0} = \frac{(1-f)i}{f}.$$

The AOQ is therefore $\frac{(1-f)}{fN_0}$ multiplied by the expected value of H and is therefore

$$\begin{aligned} \text{(7.1)} \qquad & \frac{(1-f)}{fN_0}\left[M^* - \sum_{i=0}^{M^*-1} (M^* - i)\binom{N^*}{i} p^i (1-p)^{N^*-i}\right] \\ & \qquad = L\left[1 - \frac{1}{M^*}\sum_{i=0}^{M^*-1} (M^* - i)\binom{N^*}{i} p^i (1-p)^{N^*-i}\right]. \end{aligned}$$

The reduction from the original quality p to the AOQ was achieved by inspecting a fraction of units which is $\frac{1}{p}$ times the reduction in the frequency of defectives. Hence, with probability one, the fraction of units inspected when the production process is in statistical control is

$$\text{(7.2)} \qquad I = 1 - \frac{L}{p} + \frac{(1-f)}{pN^*}\sum_{i=0}^{M^*-1} (M^* - i)\binom{N^*}{i} p^i (1-p)^{N^*-i}.$$

When $p \geq \frac{L}{1-f}$ we see from Section 4 that the third term of the right member of (7.2) represents the price paid in fraction of inspection above the minimum in return for the local stability achieved. When $p < \frac{L}{1-f}$ the additional inspection is of course $I - f$.

As N_0 becomes larger, SPC becomes more and more like SPA, and consequently the amount of inspection tends to the minimum. As N_0 becomes smaller, the degree of local stability achieved becomes higher and must be paid for by an increasing amount of inspection. An illustrative example will be given in the next section. It has already been pointed out that the mere existence of statistical control implies a considerable amount of local stability even when SPA is applied.

The only practical difficulty which may arise in evaluating the formulas in (7.1) and (7.2) might come from attempting to evaluate

$$T' = \sum_{i=0}^{M^*-1} (M^* - i) \binom{N^*}{i} p^i (1 - p)^{N^*-i}.$$

For those values of the parameters which are likely to occur in application, a good approximation to T' (exactly how good we shall not investigate here) is given by

$$T = \sum_{i=0}^{M^*-1} (M^* - i) \frac{e^{-N^*p}(N^*p)^i}{i!}.$$

A table of T for integral values of M^* from 2 to 16 and for integral values of N^*p from 1 to 25 is given below. The computations were performed under the direction of Mr. Mortimer Spiegelman of the Metropolitan Life Insurance Company, to whom the authors are deeply obliged.

Table of $T = \sum_{i=0}^{M^*-1} (M^* - i) \frac{e^{-N^*p}(N^*p)^i}{i!}$

$M^* - 1$	N^*p											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1.10	.54	.25	.11	.05	.02	.01	.00	.00	.00	.00	.00
2	2.02	1.22	.67	.35	.17	.08	.04	.02	.01	.00	.00	.00
3	3.00	2.08	1.32	.78	.44	.23	.12	.06	.03	.01	.01	.00
4	4.00	3.02	2.13	1.41	.88	.52	.29	.16	.08	.04	.02	.01
5	5.00	4.01	3.05	2.20	1.49	.96	.59	.35	.20	.11	.06	.03
6	6.00	5.00	4.02	3.08	2.26	1.57	1.04	.66	.41	.24	.14	.08
7	7.00	6.00	5.01	4.03	3.12	2.31	1.64	1.12	.73	.46	.28	.17
8	8.00	7.00	6.00	5.01	4.05	3.16	2.37	1.71	1.19	.79	.51	.32
9	9.00	8.00	7.00	6.00	5.02	4.08	3.20	2.43	1.77	1.25	.85	.56
10	10.00	9.00	8.00	7.00	6.01	5.03	4.10	3.24	2.48	1.83	1.31	.91
11	11.00	10.00	9.00	8.00	7.00	6.01	5.05	4.13	3.28	2.53	1.89	1.37
12	12.00	11.00	10.00	9.00	8.00	7.01	6.02	5.07	4.16	3.32	2.58	1.95
13	13.00	12.00	11.00	10.00	9.00	8.00	7.01	6.03	5.08	4.19	3.36	2.63
14	14.00	13.00	12.00	11.00	10.00	9.00	8.00	7.01	6.04	5.10	4.22	3.40
15	15.00	14.00	13.00	12.00	11.00	10.00	9.00	8.01	7.02	6.05	5.12	4.25

8. The SP of H. F. Dodge. H. F. Dodge [1] has proposed a very interesting SP for continuous production. The plan is defined by two constants i and f and may be described as follows: Begin with complete inspection of the units consecutively as produced and continue such inspection until i units in succession are found non-defective. Thereafter inspect a fraction f of the units. Continue partial inspection until a defect is found. Then start complete inspection again and continue until i units in succession are found non-defective. Repeat the procedure.

Dodge [1] derived formulas for determining the AOQL corresponding to any

pair i and f, under the assumption that the production process is in a state of statistical control. Dodge's formulas for the AOQL are not necessarily valid if we do not make this restriction on the production process, i.e., if we admit that the probability p that a unit will be defective may vary in any arbitrary way during the production process. This, of course, is not a criticism of the derivation of the formulas; it cannot be considered surprising that a formula is not valid under assumptions different from those under which it was derived. However, it is relevant to point out the fact that the Dodge SP does not guarantee the AOQL under all circumstances, so that care must be taken to ensure that certain requirements are met. Exactly what these requirements are is not known; statistical control is a sufficient condition, but is probably not necessary and could be weakened. It seems likely to the authors that, if p varies only slowly (with N) with infrequent "jumps," the Dodge SP will produce results which will exceed the AOQL by little, if at all. But if the "jumps" are numer-

Table of $T = \sum_{i=0}^{M^*-1} (M^* - i) \frac{e^{-N^*p}(N^*p)^i}{i!}$

(Continued)

$M^* - 1$	N^*p												
	13	14	15	16	17	18	19	20	21	22	23	24	25
1	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
2	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
3	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
4	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
5	.02	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
6	.04	.02	.01	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00
7	.10	.05	.03	.02	.01	.00	.00	.00	.00	.00	.00	.00	.00
8	.20	.12	.07	.04	.02	.01	.01	.00	.00	.00	.00	.00	.00
9	.36	.23	.14	.08	.05	.03	.01	.01	.00	.00	.00	.00	.00
10	.61	.40	.26	.16	.10	.06	.03	.02	.01	.01	.00	.00	.00
11	.97	.66	.44	.29	.18	.11	.07	.04	.02	.01	.01	.00	.00
12	1.43	1.02	.71	.48	.32	.20	.13	.08	.05	.03	.02	.01	.01
13	2.00	1.48	1.07	.75	.52	.35	.23	.15	.09	.06	.03	.02	.01
14	2.68	2.05	1.54	1.12	.80	.55	.38	.25	.16	.10	.07	.04	.02
15	3.44	2.72	2.10	1.59	1.17	.84	.59	.41	.27	.18	.12	.07	.05

ous and appropriately spaced it is possible to exceed the AOQL by substantial amounts, as the example below will show. The Dodge plan was intended to serve as an aid to the detection and correction of malfunctioning of the production process and this use would tend to prevent the occurrence of such a phenomenon. Parenthetically, it should be remarked that the information obtained in the course of inspection according to either the plans discussed in this paper or any reasonable scheme should, if possible, be sent at once to the producing divisions for their guidance.

An example to show that the AOQL can be exceeded can be constructed as

follows: Let $i = 54$ and $f = 0.1$. Then according to the graphs of [1], page 272, the AOQL should be 0.02. Define a sequence of 60 successive units free of defectives as a segment of type 1, and a sequence of 60 successive units where the production process is in statistical control with $p = 0.1$, as a segment of type 2. Suppose that the sequence of units produced consists of segments of types 1 and 2 always alternating. Then it follows that the first item inspected in a segment of type 2 is always inspected on a partial inspection basis. We now assume that, unless the occurrence of a defective has previously terminated partial inspection, the 1st, 11th, 21st, 31st, 41st, and 51st items in a segment of type 2 will be chosen for partial inspection, and if the 1st item is found defective, the entire segment of type 2 will be cleared of defectives. (Both of these assumptions favor the Dodge SP.) Then the situation is as described in the following table:

	(1) *Probability of first terminating partial inspection at each item*	(2) *Expected number of defectives remaining in segment of type 2 after partial inspection has been terminated*	(3) (1) x (2)
1st	.1	0	0
11th	$(.9)(.1) = .09$	.9	.081
21st	$(.9)^2(.1) = .081$	1.8	.1458
31st	$(.9)^3(.1) = .0729$	2.7	.19683
41st	$(.9)^4(.1) = .06561$	3.6	.236196
51st	$(.9)^5(.1) = .059049$	4.5	.2657205

Probability that an entire segment of type 2 will be partially inspected	*Expected number of defectives left in a segment of type 2 which has been inspected only partially*	*Product*
$(.9)^6 = .531441$	5.4	2.8697814
		Sum = 3.7953279

The AOQ is therefore $\dfrac{3.7953279}{120} = .0316+$, while $L = .02$.

It is therefore difficult to compare the Dodge plan with any of the plans described in this paper with respect to their effect on a production process not in statistical control. If the production process is in statistical control, then, as we have already seen, SPA requires minimum inspection (and, incidentally, because of the existence of statistical control, produces a fair degree of local stability). If, when statistical control exists, one requires both maintenance of a given AOQL and a higher degree of local stability than is produced by SPA, the relevant comparison is between the Dodge plan and SPC. Both will probably give good results as regards local stability, but it is not possible at present to make

these intuitive notions precise, as we have not given an exact definition of local stability. The following example (in which statistical control is assumed) may not be unrepresentative of what the situation is with regard to the amount of inspection required.

Fraction of product inspected under the Dodge plan and under SPC when
$L = .045$ $f = .1$

p	Fraction of product inspected under the Dodge plan	Fraction of product inspected under SPC when		
		$N_0 = 400$	$N_0 = 1000$	$N_0 = 2000$
.01	.12	.12	.10	.10
.02	.15	.17	.11	.10
.03	.19	.22	.14	.11
.04	.23	.28	.19	.15
.05	.28	.34	.26	.21
.06	.33	.40	.33	.29
.07	.39	.45	.39	.37
.08	.45	.50	.46	.44
.09	.52	.54	.51	.50
.10	.58	.57	.55	.55

The decrease in inspection required by SPC as N_0 increases is evident in this table. When $N_0 = 2000$ SPC requires less inspection than the Dodge plan, when $N_0 = 400$ it requires more inspection than the Dodge plan. How the various degrees of local stability achieved compare remains an open question. The case when $N_0 = 400$ probably lies in the region where SPC is inefficient (as regards amount of inspection) and corresponds to a high degree of local stability.

We note that both plans call for increased inspection as the quality worsens (p increases). If the manufacturer is required to pay for the inspection this serves as an added incentive to improve quality of output.

REFERENCES

[1] H. F. Dodge, *Annals of Math. Stat.*, 14 (1943), p. 264.
[2] A. Kolmogoroff, *Comptes Rendus Acad. Sciences*, 191 (1930), p. 910.
[3] Maurice Fréchet, *Généralités sur les Probabilités. Variables aléatoires.* Paris, 1937.
[4] H. F. Dodge and H. G. Romig, *Bell System Tech. Journal*, 20 (1941), p. 1.

THE EXPECTED VALUE AND VARIANCE OF THE RECIPROCAL AND OTHER NEGATIVE POWERS OF A POSITIVE BERNOULLIAN VARIATE[1]

By Frederick F. Stephan

War Production Board, Washington

1. Introduction. The expected value of the reciprocal of a Bernoullian variate appears in certain problems of random sampling wherein both practical considerations and mathematical necessity make zero an inadmissible value of the variate. This special condition excluding zero is necessary from a practical standpoint because statistics can not be calculated from an empty class. It is a necessary condition, in the mathematical sense, for the expected value, and variances involving it, to be finite. When subject to this condition the Bernoullian variate will be designated the *positive* Bernoullian variate.

There appears to be no simple expression for the expected value of the reciprocal such as there is for the expected value of positive integral powers of the positive Bernoullian variate. This paper presents in (15) a factorial series, which can be computed conveniently to any desired number of terms by means of the recursion relation (18). Upper and lower bounds on the remainder may be computed readily from (20), (21), (23), (24), and (26) and the approximation may be improved by adding an estimate of the remainder taken between these bounds. A factorial series for the expected value of negative integral powers is given in (34). A factorial series for the expected value of the reciprocal of the positive hypergeometric variate is given in (53). Series for the variances follow directly from the series for expected values.

A simple example of the sampling problems in which this expected value appears is presented by the following instance of estimates derived from samples of variable size:

An infinite population consists of items of two kinds or classes, A and B. Lots of N items each are drawn at random. In such lots the number of items, x', that are of class A is an ordinary Bernoullian variate. Next, every lot composed entirely of items of class B is discarded. This excludes all lots for which $x' = 0$. From each remaining lot the $N - x'$ items of class B are set aside, leaving a sample composed entirely of items of class A. The number of such items, x, varies from sample to sample. It will be designated a positive Bernoullian variate since $x = x'$ if $x' > 0$ and x does not exist if $x' \leq 0$. Finally, let there be associated with each item in class A a particular value of a variable, y, the variance of which in A is σ^2. Then if the mean value of y is computed for each sample, the error variance of such means is $E(\sigma^2/x) = \sigma^2 E(1/x)$.

Instances similar to that just described occur in the design of sampling surveys from which statistics are to be obtained separately for each of several classes

[1] Developed from a section of a paper presented to the Washington meeting of the Institute of Mathematical Statistics on June 18, 1943.

of the population, i.e., each statistic is to be computed from some part of the sample instead of all of it. They also occur in certain sampling problems in which some of the items drawn for a sample turn out to be blanks.

A related problem concerning the error variance of the proportion of males among infants born in any one year was considered by G. Bohlmann in a paper on approximations to the expected value and standard error of a function [1]. His approach to the problem was to expand the function in a Taylor series and take the expected value of each term. The conditions under which the resulting series converges were developed for certain functions of a Bernoullian variate. The present paper provides a different and, in certain respects, superior approach to the problem employing a method due to Stirling [2]. While the method is applied to the reciprocal and negative powers it is also applicable to certain other functions of a Bernoullian variate.

2. The positive Bernoullian variate. Let x be a random variate defined by a Bernoullian probability function subject to the special condition $x > 0$. The probability of x in n is

$$P(x) = \binom{n}{x} p^x q^{n-x}/(1 - q^n) \tag{1}$$

where x and n are integers, $1 \leq x \leq n$, and

$$\binom{n}{x} = \frac{n!}{x!(n - x)!}. \tag{2}$$

The probabilities p and q are constants, $0 < p = 1 - q < 1$.

The divisor $1 - q^n$ arises from the condition excluding zero. (Bohlmann omits this factor, assuming that q^n is negligible, an assumption that is not always valid. In fact, $q^n \sim e^{-np}$.) An extension of this condition to exclude all values of x less than a specified constant will be considered in a later section.

Throughout this paper summation is understood to be from $x = 1$ to $x = n$ unless it is shown otherwise.

3. Expected values and moments. The expected values of x and its positive integral powers are

$$E(x) = np/(1 - q^n) \tag{3}$$

$$E(x^2) = (npq + n^2p^2)/(1 - q^n) \tag{4}$$

and, in general

$$E(x^i) = \nu_i/(1 - q^n) = \sum_j \mathfrak{S}_i^j j! \binom{n}{j} \frac{p^j}{1 - q^n}, \qquad i > 0 \tag{5}$$

where ν_i is the ith moment about zero of an ordinary Bernoullian variate with the same n and p and the $\mathfrak{S}_i^j$ are the Stirling numbers of the second kind (see Table 1).

The moments about $E(x)$ are somewhat more complicated than the corre-

sponding moments of the ordinary Bernoullian variate. For example, the variance

$$E\{(x - E(x))^2\} = \frac{npq}{1 - q^n} - \frac{n^2p^2q^n}{(1 - q^n)^2} \tag{6}$$

and the third moment

$$E\{(x - E(x))^3\} = \frac{(q - p)npq}{1 - q^n} - \frac{3n^2p^2q^{n+1}}{(1 - q^n)^2} + \frac{n^3p^3q^n(1 + q^n)}{(1 - q^n)^3}. \tag{7}$$

The moments about np, the first moment of an ordinary Bernoullian variate, are

$$E\{(x - np)^i\} = (\mu_i + (-1)^{i-1}(np)^iq^n)/(1 - q^n) \tag{8}$$

TABLE 1

Stirling numbers of the second kind, $\mathfrak{S}_i^j$

$i \backslash j$	1	2	3	4	5	6
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	1	3	1	0	0	0
4	1	7	6	1	0	0
5	1	15	25	10	1	0
6	1	31	90	65	15	1
7	1	63	301	350	140	21
8	1	127	966	1,709	1,050	266
9	1	255	3,025	7,770	6,951	2,646
10	1	511	9,330	34,105	42,525	22,827

where μ_i is the ith moment, about the mean, of an ordinary Bernoullian variate with the same values n and p.

The expected value of the reciprocal is

$$E\left(\frac{1}{x}\right) = \frac{1}{1 - q^n}\left\{\frac{1}{1}npq^{n-1} + \frac{1}{2}\cdot\frac{1}{2}n(n-1)p^2q^{n-2} + \cdots + \frac{1}{i}\binom{n}{i}p^iq^{n-i} + \cdots + \frac{1}{n}p^n\right\}. \tag{9}$$

This equation is not suitable for the computation of $E(1/x)$ to a satisfactory degree of approximation unless np is small, say less than 5 for most purposes. The number of terms necessary to obtain a computed value with four significant figures, for example, may be estimated to be approximately $8\sqrt{npq/(- q^n)}$.

Expressed as a function of q, $E(1/x)$ becomes

$$E\left(\frac{1}{x}\right) = \frac{1}{1 - q^n}\,\Sigma\,\frac{q^{x-1} - q^n}{n - x + 1} \tag{10}$$

a series which may be convenient for small values of q.

$E(1/x)$ may be expanded in a power series by Taylor's Theorem. It may

also be expanded in a finite series of expected values of powers, either in $E(x)$, $E(x^2)$, $\cdots$ or in $E(x - c)$, $E(x - c)^2$, $\cdots$ c being any positive constant. The second of these three series may be obtained by expanding $\frac{1}{x}\left(1 - \frac{x}{c}\right)^t$ and taking expected values, and the third by dividing out $\frac{1}{x} = \frac{1}{c + (x - c)}$ and taking expected values. For all three expansions, however, the terms become progressively more complicated and laborious to compute. A simpler and more convenient series for actual computations may be obtained by expanding $1/x$ in a factorial series.

4. Expansion of $E(1/x)$ in a series of inverse factorials. It is easy to prove by induction that, $x > 0$,

$$\frac{1}{x} = \frac{0!}{x+1} + \frac{1!}{(x+1)(x+2)} + \cdots + \frac{(i-1)!x!}{(x+i)!} + \cdots + \frac{(t-1)!x!}{(x+t)!} + R_t(x) \tag{11}$$

where

$$R_t(x) = t!(x-1)!/(x+t)! \tag{12}$$

is the remainder after the first t terms. This is, of course, an expansion in Beta functions. It is also a simple special case of the expansion of a function in a "faculty series" or series of inverse factorials [3] with an exact expression for the remainder.

Let

$$s_i = \Sigma \binom{n+i}{x+i} \frac{p^{x+i} q^{n-x}}{1-q^n} = \frac{1}{1-q^n}\left(1 - \sum_{x=0}^{i} \binom{n+i}{x} p^x q^{n+i-x}\right). \tag{13}$$

Then, since

$$\Sigma \frac{x!}{(x+1)!} \binom{n}{x} p^x q^{n-x} = \frac{n!\, s_i (1 - q^n)}{(n+i)!\, p^i} \tag{14}$$

the expected value of (11) is

$$E\left(\frac{1}{x}\right) = \frac{0!\, s_i}{(n+1)p} + \frac{1!\, s_2}{(n+1)(n+2)p^2} + \cdots + \frac{(i-1)!\, n!\, s_i}{(n+1)!\, p^i} + \cdots + \frac{(t-1)!\, n!\, s_t}{(n+t)!\, p^t} + \Sigma R_t(x) P(x). \tag{15}$$

When developed as infinite series, both (11) and (15) are convergent since the remainders $R_t(x) \to 0$ as $t \to \infty$.

For computing purposes it is convenient to write

$$E\left(\frac{1}{x}\right) = \sum_{i=1}^{t} u_i + E(R_t(x)) \tag{16}$$

in which, since

$$s_i = s_{i-1} - q\binom{n+i-1}{i} \frac{p^i q^{n-1}}{1-q^n}, \tag{17}$$

the following recursion relation exists between u_i and u_{i-1}

$$(18)\qquad u_i = \frac{(i-1)!\,n!\,s_i}{(n+i)!\,p^i} = \frac{(i-1)u_{i-1} - k/i}{(n+i)p}, \qquad i > 1; \qquad u_1 = \frac{1-k}{(n+1)p}$$

where

$$(19)\qquad k = npq^n/(1-q^n) \sim np/(e^{np}-1).$$

This reduces the computing of the u_i to a simple repetitive procedure. The computing is still simpler in those problems in which, for the degree of precision desired, k is negligible.

An estimate of $E(R_t(x))$ should be added to the sum in (16) to improve the approximation. To determine a suitable estimate, a lower bound for the expected value of the remainders may be computed from one of the following inequalities:

$$(20)\qquad \begin{aligned} E(R_t(x)) &= \Sigma \frac{t}{x}\,\frac{(t-1)!\,x!}{(x+t)!}\,P(x) \\ &= \Sigma t\left(\frac{1}{m} - \frac{x-m}{m^2} + \frac{(x-m)^2}{m^2 x}\right)\frac{(t-1)!\,x!}{(x+t)!}\,P(x) \\ &> \frac{1}{m}\,tu_t - \frac{1}{m^2}\,t(t-1)u_{t-1} + \frac{m+t}{m^2}\,tu_t, \qquad m \neq 0 \end{aligned}$$

which is maximized by setting $m = \{(t-1)u_{t-1} - tu_t\}/u_t$, whence

$$(21)\qquad E(R_t(x)) > tu_t^2/\{(t-1)u_{t-1} - tu_t\}, \qquad t > 1.$$

Also, since when $m = E(x)$

$$(22)\qquad \Sigma(x-m)\,\frac{(t-1)!\,x!}{(x+t)!}\,P(x) < \Sigma(x-m)P(x) = 0,$$

a simpler inequality is

$$(23)\qquad E(R_t(x)) > tu_t(1-q^n)/np.$$

Further, if only the first $c < n$ terms in (20) are taken,

$$(24)\qquad E(R_t(x)) > \sum_{x=1}^{c} \frac{t!\,(x-1)!}{(x+t)!}\,P(x) = \sum_{x=1}^{c} v_x$$

where

$$(25)\qquad v_1 = \frac{k}{(t+1)q} \quad \text{and} \quad v_x = \frac{(x-1)(n-x+1)p}{x(x+t)q}\,v_{x-1}.$$

An upper bound may be computed from

$$(26)\qquad E(R_t(x)) < \begin{cases} tu_t & (26.1) \\ \frac{1}{2}\,tu_t + \frac{1}{2}\,v_1 & (26.2) \\ \frac{1}{3}\,tu_t + \frac{2}{3}\,v_1 + \frac{1}{6}\,v_2 & (26.3) \\ \quad\cdots & \\ \frac{1}{j}\,tu_t + \sum_{x=1}^{j-1}\left(\frac{1}{x} - \frac{1}{j}\right)v_x & (26.j) \end{cases}$$

the choice among which may be governed by computing convenience. Taken with (16), these inequalities provide lower and upper bounds for $E(1/x)$.

5. Examples. Two examples will serve to illustrate the factorial series (15).

EXAMPLE 1

Computation of $E(1/x)$ for $n = 100$ and $p = 0.1$

$np = 10 \qquad k = .000{,}265{,}621 \qquad E(1) = .111{,}527$

t	*Binomial sum of t terms*	*Sum of t terms*	*Factorial series lower bounds**		*Upper bound***
1	.000,295	.098,984	.099,647		.132,167
2	.001,107	.108,675	.109,006	(.111,034)	.115,247
3	.003,071	.110,548	.110,752	(.111,313)	.112,498
4	.007,039	.111,082	.111,223	(.111,381)	.111,852
5	.013,813	.111,280	.111,385	(.111,452)	.111,657
6	.023,743	.111,370	.111,452	(.111,478)	.111,587
7	.036,442	.111,416	.111,483	(.111,489)	.111,556
8	.050,796	.111,444	.111,500	(.111,497)	.111,544
9	.065,287	.111,461	.111,509	(.111,503)	.111,537
10	.078,474	.111,472	.111,514	(.111,508)	.111,534
11	.089,372	.111,481	.111,518	(.111,511)	.111,532
12	.097,604	.111,487	.111,520		.111,530
13	.103,320	.111,492	.111,521		.111,529
14	.106,985	.111,495	.111,523		.111,529
15	.109,164	.111,498	.111,524		.111,529
16	.110,369	.111,501	.111,524		.111,528
17	.110,992	.111,503	.111,525		.111,528
18	.111,294	.111,505	.111,525,4		.111,527,5
19	.111,431	.111,506	.111,525,6		.111,527,3
20	.111,489	.111,508	.111,525,8		.111,527,1
...					
24	.111,526				
...					
100	.111,527 (end of series)				

* Sum of t terms plus lower bound for $E(R(x))$ from (24) with $c = 3$. Numbers in parentheses are calculated from (21).

** Sum of t terms plus upper bound on $E(R(x))$ from (26.3).

EXAMPLE 2

Computation of $E(1/x)$ for $n = 1000$ and $p = 0.3$

$np = 300 \qquad k = 9.7 \times 10^{-14}$

t	*Sum of t terms*	*Factorial series upper and lower bounds**
1	.003,330,003,330	
2	.003,341,081,185	$\begin{cases} .003{,}346{,}7 \\ .003{,}341{,}0 \ (.003{,}341{,}155{,}4) \end{cases}$

* Computed as in Example 1.

t	*Sum of t terms*	*Factorial series upper and lower bounds**
3	.003,341,154,817	.003,341,211 .003,341,155
4	.003,341,155,549	.003,341,156,29 .003,341,155,56
5	.003,341,155,559	.003,341,155,58 .003,341,155,57

For the binomial series, the sum of the largest eight terms of (9), not the first eight terms, is approximately .0007 which is less than 1/4 of the value of $E(1/x)$.

In the first example the value of np is almost small enough to make computation by (9) convenient. In the second example about 120 terms of (9) must be computed to obtain an approximation to four significant figures but only four terms of the factorial series are needed to obtain seven significant figures. It is evident that as np increases, the number of terms of (16) required to obtain an approximation to a given number of significant figures decreases. The opposite is true of (9) as n increases, or as p approaches a value near 1/2.

6. Extending the special condition. In some sampling problems all values of x less than a specified value, g, and greater than another specified value, h, are inadmissible. Then the probability of x in n is

$$P(x \mid g, h) = \binom{n}{x} p^x q^{n-x}/s_{0,g,h}, \qquad g \leq x \leq h, \tag{27}$$

where

$$s_{0,g,h} = \sum_{x=g}^{h} \binom{n}{x} p^x q^{n-x}. \tag{28}$$

With this new condition, $E(1/x)$ is given by (15) if s_i is replaced by

$$s_{i,g,h} = \sum_{x=g}^{h} \binom{n+i}{x+i} \frac{p^{x+i} q^{n-x}}{s_{0,g,h}} \tag{29}$$

and the summation in the remainder term is from g to h. Also since

$$s_{i,g,h} = s_{i-1,g,h} - \frac{q}{s_{0,g,h}} \left\{ \binom{n+i-1}{g+i-1} p^{g+i-1} q^{n-g} - \binom{n+i-1}{h+i} p^{h+i} q^{n-h-1} \right\} \tag{30}$$

a recursion relation similar to (18) may be used in computing

$$\begin{aligned} u_{i,g,h} &= \frac{(i-1)!\,n!\,s_{i,g,h}}{(n+i)!\,p^i} \\ &= \frac{(i-1)u_{i-1,g,h} - (i-1)!\,\{k_g/(g+i-1)! + k_{h+1}/(h+i)\}}{(n+i)p} \end{aligned} \tag{31}$$

where

$$k_g = \frac{n!\,p^g\,q^{n-g+1}}{(n-g)!\,s_{0,g,h}} \tag{32}$$

$$k_h = \frac{n!\,p^h\,q^{n-h+1}}{(n-h)!\,s_{0,g,h}}. \tag{33}$$

The inequalities (20) to (23) inclusive and (26) are applicable to this extension on substitution of $u_{i,g,h}$ for u_i.

7. Expansion of $E(x^{-a})$ in a factorial series. Equation (11) may be extended to other negative integral powers of x. If a is a positive integer

$$\begin{aligned} E(x^{-a}) = \Sigma \frac{1}{x^a} P(x) &= \frac{b_{1,a}\,s_1}{(n+1)p} + \frac{b_{2,a}\,s_2}{(n+1)(n+2)p^2} \\ &\quad + \cdots + \frac{b_{t,a}\,s_t\,n!}{(n+t)!\,p^t} + \Sigma R_t'(x)P(x) \end{aligned} \tag{34}$$

where

$$R_t'(x) = \sum_{j=1}^{a} b_{t+1,j} \frac{x^{j-1}\,x!\,P(x)}{(x+t)!\,x^a} \tag{35}$$

and the $b_{i,j}$ are the absolute values of the Stirling numbers of the first kind (see Table 2) formed by the recursion relation

$$b_{i,j} = b_{i-1,j-1} + (i-1)b_{i-1,j}, \qquad b_{i,j} = 0 \quad \text{if} \quad j > i \quad \text{or} \quad j < 1. \tag{36}$$

It is evident that

$$\sum_{j=1}^{i} b_{i,j} = i! \tag{37}$$

$$b_{i,1} = (i-1)! \quad \text{and} \quad b_{i,j} < i! \quad \text{if} \quad j > 1, \tag{38}$$

whence

$$\begin{aligned} R_t'(1) &= \frac{1}{t+1} P(1) \\ R_t'(x) &< \frac{((t+1)! - t!)x!P(x)}{2(x+t)!}, \qquad x > 1 \\ &< \frac{1}{(t+1)} P(x). \end{aligned} \tag{39}$$

Hence $R_t'(x) \to 0$ and $E(R_t'(x)) \to 0$ as $t \to \infty$ and the sum of the first t terms of (34) converges to $E(x^{-a})$ as $t \to \infty$.

The following recursion relation corresponding to (18) provides a simple procedure for computing:

$$u_{i,a} = b_{i,a}u_i/(i-1)! = b_{i,a}\frac{(u_{i-1,a}/b_{i-1,a}) - k/i!}{(n+1)p}. \tag{40}$$

The computing procedure, then, follows a cycle of four simple operations:
1. Divide $\{k/(i-1)!\}$ by i.
2. Subtract the quotient from $\{u_{i-1,a}/b_{i-1,a}\}$.
3. Divide the difference by $\{(n+i+1)p\} + p$. The quotient is $u_{i,a}/b_{i,a}$.
4. Multiply this quotient by $b_{i,a}$.

TABLE 2

*Absolute values of Stirling numbers of the first kind, $b_{i,j}$**

i \ j	1	2	3	4	5	6
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	2	3	1	0	0	0
4	6	11	6	1	0	0
5	24	50	35	10	1	0
6	120	274	225	85	15	1
7	720	1,764	1,624	735	175	21
8	5,040	13,068	13,132	6,769	1,960	322
9	40,320	109,584	118,124	67,284	22,449	4,536
10	362,880	1,026,576	1,172,700	723,680	269,325	63,273

* These numbers are also known as differential coefficients of zero [4].

The expressions in braces are quantities obtained in the preceding cycle.

The $u_{i,a}$ may also be calculated from (18), or checked by such a calculation.

A lower bound for $E(R'(x))$ after t terms may be calculated from the first c terms of

$$\begin{aligned} E(R'(x)) &= \sum R_t'(x)P(x) > \sum_{x=1}^{c} R_t'(x)P(x) \\ &= \sum_{x=1}^{c}\sum_{j=1}^{a}\frac{b_{t+1,j}\,n!\,p^x q^{n-x}}{x^{a-j+1}(x+t)!\,(n-x)!\,(1-q^n)} \end{aligned} \tag{41}$$

or from an inequality similar to (23)

$$E(R'(x)) > \frac{u_t}{(t-1)!}\sum_{j=1}^{a}\frac{b_{t+1,j}}{(E(x))^{a-j+1}} \tag{42}$$

which may also be written

$$E(R'(x)) > \frac{u_t}{(t-1)!\,(E(x))^{a+1}}\Big\{(E(x)+t)(E(x)+t-1)\cdots E(x) \tag{43}$$
$$- \sum_{j=a+1}^{t+1} b_{t+1,j}(E(x))^j\Big\}.$$

An upper bound may be calculated from

$$E(R'(x)) < \frac{u_t}{(t-1)!}\sum_{j=1}^{a} b_{t+1,j} < t(t+1)u_t \tag{44}$$

or

$$E(R'(x)) < \sum_{x=1}^{c} R'(x)P(x) + \sum_{x=c+1}^{n}\sum_{j=1}^{a} b_{t+1,j}\,\frac{x!P(x)}{(x+t)!c^{a-j+1}}$$
$$< \sum_{x=1}^{c} R'(x)P(x) + \frac{u_t}{(t-1)!}\sum_{j=1}^{a}\frac{b_{t+1,j}}{c^{a-j+1}} = \sum_{x=1}^{c} R'(x)P(x) \tag{45}$$
$$+ \frac{u_t}{(t-1)!c^{a+1}}\Big\{(c+t)(c+t-1)\cdots c - \sum_{j=a+1}^{t+1} b_{t+1,j}c^j\Big\}.$$

8. The positive hypergeometric variate. The theory of sampling without replacement from a finite population rests on the hypergeometric variate. Its probability function is

$$P(x \mid N, M, n) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}. \tag{46}$$

In applications to finite sampling, N is the number of items in the population, M is the number of them that are of a certain kind, n is the number of items drawn for the sample, and x is the number of items of the designated kind in the sample.

As in the case of the Bernoullian variate, it is necessary to exclude zero in defining the expected value of $1/x$. The probability function of the positive hypergeometric variate, then, is

$$P_H(x) = P(x \mid N, M, n)/s_0, \qquad x > 0 \tag{47}$$

where

$$s_0 = 1 - P(0 \mid N, M, n). \tag{48}$$

Throughout this section the notation will have reference to (47) instead of (1). The expected values of positive integral powers of x are

$$E(x) = Mn/(Ns_0) \tag{49}$$

$$E(x^2) = \frac{1}{s_0}\left\{\frac{M(M-1)n(n-1)}{N(N-1)} + \frac{Mn}{N}\right\} \tag{50}$$

and, in general,

$$E(x^i) = \sum_{j=1}^{i} \mathfrak{S}_i^j E(x!/(x-j)!) \tag{51}$$

where the $\mathfrak{S}_i^j$ are the Stirling numbers of the second kind and

$$E\left(\frac{x!}{(x-j)!}\right) = \frac{M!n!(N-j)!}{(M-j)!(n-j)!N!s_0}. \tag{52}$$

The factorial series corresponding to (16) is

$$E\left(\frac{1}{x}\right) = \sum \frac{1}{x} P_H(x) = \sum_{i=1}^{t} u_i + E(R_t(x)) \tag{53}$$

where

$$u_i = \Sigma \frac{(i-1)!x!}{(x+i)!} P_H(x) \tag{54}$$

and

$$E(R_t(x)) = \Sigma \frac{t!(x-1)!}{(x+t)!} P_H(x). \tag{55}$$

The u_i may be computed from

$$\begin{aligned} u_1 &= \frac{(N+1)s_1}{(M+1)(n+1)s_0} \\ &= \frac{1}{s_0}\left\{\frac{N+1}{(M+1)(n+1)} - \frac{(N-M)!(N-n)!}{N!(N-M-n-1)!(M+1)(n+1)}\right\} \end{aligned} \tag{56}$$

and the recursion relation

$$u_i = \frac{(N+i)s_i}{(M+i)(n+i)s_{i-1}} u_{i-1} \tag{57}$$

where

$$s_i = 1 - \sum_{x=0}^{i} P(x \mid N+i, M+i, n+i). \tag{58}$$

The computing is quite simple in those instances in which $1 - s_t$ is negligible.

Corresponding to (26), an upper bound for the expected value of the remainders after t terms may be computed from

$$E(R_t(x)) < \begin{cases} tu_t & (59.1) \\ \frac{1}{2}tu_t + \frac{1}{2}P_H(1)/(t+1) & (59.2) \\ \frac{1}{3}tu_t + \frac{2}{3}\frac{P_H(1)}{t+1} + \frac{1}{6}\frac{P_H(2)}{(t+1)(t+2)} & (59.3) \\ \cdots & \\ \frac{1}{j}tu_t + t!\sum_{x=1}^{j-1}\left(\frac{1}{x} - \frac{1}{j}\right) P_H(x) \frac{x!}{(x+t)!}. & (59.j) \end{cases} \tag{59}$$

A lower bound for the expected value of the remainders may be computed from one of the following inequalities corresponding to (23), (21) and (24)

$$E(R_t(x)) > tu_tNs_0/(Mn) \tag{60}$$

$$E(R_t(x)) > tu_t^2/\{(t-1)u_{t-1} - tu_t\} \tag{61}$$

$$E(R_t(x)) > \sum_{x=1}^{j} \frac{t!(x-1)!}{(x+t)!} P_H(x). \tag{62}$$

The expected values of other negative integral powers of the positive hypergeometric variate may be calculated from

$$E(x^{-a}) = \sum_{i=1}^{t} b_{i,a}u_i/(i-1)! + E(R_t'(x)) \tag{63}$$

where

$$R_t'(x) = \sum_{j=1}^{a} b_{t+1,j} \frac{x^{j-1}x!P_H(x)}{x^a(x+t)!}. \tag{64}$$

With $P_H(x)$ substituted for $P(x)$, (39), (42), (43), (44), and (45) provide lower and upper bounds for $E(R_t'(x))$ for the positive hypergeometric variate. Also, corresponding to (41)

$$E(R_j'(x)) > \sum_{x=1}^{c} R_t'(x)P_H(x). \tag{65}$$

9. Variance and moments of $1/x$ and x^{-a}. The variance of $1/x$, which is $E(1/x^2) - (E(1/x))^2$, may be calculated from (16) and (34), with $a = 2$, for the positive Bernoullian variate, and from (53) and (63), with $a = 2$, for the positive hypergeometric variate. Likewise, the variance of x^{-a} and the moments of $1/x$ and x^{-a} about $E(1/x)$ may be computed by the usual formulae.

REFERENCES

[1] G. Bohlmann,[2] "Formulierung und begründung zweier hilfssätze der mathematische Statistik," *Math. Annalen*, 74(1913), 341–409.

[2] E. T. Whittaker and G. Robinson, *The Calculus of Observations*, London (Second Ed.) 1937, p. 368.

[3] E. T. Whittaker and G. N. Watson, *Modern Analysis*, Cambridge (Fourth Ed.) 1927, p. 142.

[4] Charles Jordan, *Calculus of Finite Differences*, Budapest, 1937.

[2] The writer is indebted to Dr. Felix Bernstein for the reference to Bohlman.

RANDOM WALK IN THE PRESENCE OF ABSORBING BARRIERS

M. Kac

Cornell University

1. Introduction. The problem of random walk (along a straight line) in the presence of absorbing barriers can be stated as follows:

A particle, starting at the origin, moves in such a way that its displacements in consecutive time intervals, each of duration Δt, can be represented by independent random variables

$$X_1, X_2, X_3, \cdots$$

Moreover, if at some time the total (cumulative) displacement becomes $>p$ ($p \geq 0$) or $< -q$ ($q \geq 0$) the particle gets absorbed. The problem is to determine the probability that "the length of life" of the particle is greater than a given number t. This problem also admits an interpretation in terms of a game of chance in which the player quits when he loses more than q or wins more than p. An interesting paper on this type of problem by A. Wald[1] appeared recently in the *Annals*. Wald assumes that the X's are identically distributed and that their mean and standard deviation are different from 0.[2] He is then mostly interested in the limiting case when both the mean and the standard deviation become small. The object of this paper is to propose a different method of attack which in some cases leads to an answer in closed form. The method we use has been employed repeatedly in statistical mechanics in the study of the so called order-disorder problem. It is due, I believe, to E. W. Montroll[3]. As far as the author knows this method was never used in connection with the classical probability theory and this seems to furnish an additional reason for publishing this paper.

2. The simplest discrete case. We assume that each X is capable of assuming the values 1 and -1 each with probability $\frac{1}{2}$, and for simplicity sake we let $\Delta t = 1$. Note that, unlike in Wald's case, the mean of X is 0. Denote by N the random variable which represents the "length of life" of the particle and let (m an integer)

$$\delta(m) = \begin{matrix} \frac{1}{2} & m = 1 \quad \text{or} \quad m = -1, \\ 0 & \text{otherwise.} \end{matrix}$$

[1] A. Wald "On cumulative sums of random variables," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 283–296.

[2] Since this was written Professor Wald informed the author that he can easily avoid the condition that the mean should be zero.

[3] See for instance E. W. Montroll, "Statistical Mechanics of nearest neighbor systems," *Jour. of Chem. Physics*, Vol. 9 (1941), pp. 706–721.

Clearly we have (throughout this section we assume that both p and q are integers)

$$\text{Prob. } \{N > n\} = \text{Prob. } \{-q \leq X_1 \leq p,\ -q \leq X_1 + X_2 \leq p, \cdots, -q \leq X_1 + \cdots + X_b \leq p\} = \Sigma\delta(m_1)\delta(m_2)\cdots\delta(m_n),$$

where the summation is extended over all integers $m_1, m_2, \cdots m_n$ for which

$$-q \leq m_1 \leq p, \quad -q \leq m_1 + m_2 \leq p, \cdots, \quad -q \leq m_1 + m_2 + \cdots + m_n \leq p.$$

Letting

$$l_j = q + m_1 + \cdots + m_j, \qquad (j = 1, 2, \cdots, n),$$

we see that

$$\text{(1)} \qquad \text{Prob } \{N > n\} = \sum_{l_1, \cdots, l_n = 0}^{p+q} \delta(l_1 - q)\delta(l_2 - l_1)\cdots\delta(l_n - l_{n-1}).$$

Let us now consider the $(p + q + 1)$ by $(p + q + 1)$ matrix

$$\text{(2)} \qquad A = ((\delta(i - k))) = \begin{Bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{Bmatrix}.$$

It is easily seen that the sum in (1) is equal to the sum of the elements in the $(q + 1)$-st column (or row) of the matrix A^n. Thus

Prob. $\{N > n\}$ = sum of the elements of the $(q + 1)$-st column of A^n.

Denote by $\lambda_1, \lambda_2, \cdots \lambda_{p+q+1}$ the eigenvalues of the matrix A and let

$$(x_1^{(j)}, x_2^{(j)}, \cdots, x_{p+q+1}^{(j)})$$

be the normalized eigenvector of A belonging to the eigenvalue λ_j. It can be shown by elementary means[4] that

$$\lambda_j = \cos\frac{\pi j}{p + q + 2}$$

[4] Matrices of type (2) have been introduced and studied in various connections. In a paper by R. P. Boas and the present author recently accepted by the *Duke Mathematical Journal* references to several authors are given. In order to find the eigenvalues and the eigenvectors of (2) it suffices to know that

$$\begin{vmatrix} 1 & a & 0 & \cdots & \\ a & 1 & a & \cdots & \\ 0 & a & 1 & a & \cdots \\ 0 & 0 & a & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = \frac{\rho_1^{m+1} - \rho_2^{m+1}}{\rho_1 - \rho_2},$$

where m is the order of the matrix ρ_1 and ρ_2 roots of the equation $\rho^2 - \rho + a^2 = 0$.

and

$$x_k^{(j)} = \frac{\sqrt{2}}{\sqrt{p+q+2}} \sin \frac{\pi j k}{p+q+2}.$$

Denoting by R the orthogonal matrix

$$\begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_{p+q+1}^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_{p+q+1}^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{(p+q+1)} & x_2^{(p+q+1)} & \cdots & x_{p+q+1}^{(p+q+1)} \end{pmatrix}$$

and by R' the transposed of R we have (since the eigenvalues of A are simple) by a well known theorem

$$A^n = R' \begin{pmatrix} \lambda_1^n & & & 0 \\ & \lambda_2^n & & \\ & & \ddots & \\ 0 & & & \lambda_{p+q+1}^n \end{pmatrix} R.$$

It thus follows by an easy computation that the sum of the elements of the $(q+1)$-st column (row) of A^n is

$$\sum_{r=1}^{p+q+1} \sum_{j=1}^{p+q+1} \lambda_j^n x_n^{(j)} x_{q+1}^{(j)} = \sum_{j=1}^{p+q+1} \lambda_j^n x_{q+1}^{(j)} \left(\sum_{r=1}^{p+q+1} x_r^{(j)} \right).$$

We have

$$\sum_{r=1}^{p+q+1} x_r^{(j)} = \frac{\sqrt{2}}{\sqrt{p+q+2}} \sum_{r=1}^{p+q+1} \sin \frac{\pi j r}{p+q+2}$$

$$= \begin{cases} 0, & j \text{ even}, \\ \dfrac{\sqrt{2}}{\sqrt{p+q+2}} \cot \dfrac{\pi j}{2(p+q+2)}, & j \text{ odd}, \end{cases}$$

and therefore[5]

Prob. $\{M > n\}$

$$= \frac{2}{p+q+2} \sum_{j=1}^{p+q+1}{}^{*} \cos^n \frac{\pi j}{p+q+2} \sin \frac{\pi j(q+1)}{p+q+2} \cot \frac{\pi j}{2(p+q+2)},$$

where the star on the summation sign indicates that only odd j's are taken under account.

The method just illustrated is quite general but in more complicated cases the job of finding the eigenvalues and eigenvectors becomes formidable.

[5] Professor Feller has called the author's attention to the fact that similar problems and formulas can be found in Chapter III of W. Burnside's *Theory of Probability* (Cambridge, 1928). He also pointed out that the problem could be treated by means of Markoff chains.

Professor G. E. Uhlenbeck has pointed out that our formula implies a known result from the theory of Brownian motion.

Consider a free Brownian particle which at $t = 0$ is at $x = x_0(x_0 > 0)$. R. Fürth[6] has shown that the probability that between t and $t + dt$ the particle will be either at $x = 0$ or at $x = d$ $(0 < x_0 < d)$ for the first time, is given by the formula

$$dt \frac{4\pi D}{d^2} \sum_{m=0}^{\infty} (2m + 1)e^{(-\pi^2 Dt/d^2)(2m+1)^2} \sin \frac{(2m + 1)\pi x_0}{d},$$

where D is the "coefficient of diffusion."

If we treat the one-dimensional Brownian motion as a random walk with steps $\pm\Delta x$, each move lasting Δt, the probability that a particle starting from x_0 will not have reached 0 or d in the time interval $(0, t)$ can be calculated by means of our formula.

We must only put $q = x_0/\Delta x$, $p = (d - x_0)/\Delta x$, $n = t/\Delta t$ and assume that as both Δx and Δt approach 0 the ratio $(\Delta x)^2/2\Delta t$ approaches the "coefficient of diffusion" D.

An elementary computation shows that in this limit the Prob. $\{N > t/\Delta t\}$ approaches

$$\frac{4}{\pi} \sum_{j=1}^{\infty}{}^{*} \frac{1}{j} e^{(-\pi^2 j^2 D/d^2)t} \sin \frac{\pi j x_0}{d}$$

and that the differential of this expression (with a minus sign) gives exactly Fürth's expression.

3. General theory in the continuous case. We now assume that the distribution function of X possesses a continuous and even density function $\rho(x)$. We have

$$\text{Prob. } \{N > n\} = \int \cdots \int_{\Omega} \rho(x_1) \cdots \rho(x_n)\, dx_1 \cdots dx_n,$$

where the region of integration Ω is defined by the inequalities

$$-q \leq x_1 \leq p, \qquad -q \leq x_1 + x_2 \leq p, \cdots, \qquad -q \leq x_1 + \cdots + x_n \leq p$$

Introducing the new variables

$$y_j = q + x_1 + \cdots + x_j, \qquad (j = 1, 2, \cdots, n),$$

we see that the Jacobian of the transformation is 1 and

$$\begin{aligned} &\text{Prob. } \{N > n\} \\ (3) \qquad &= \int_0^{p+q} \cdots \int_0^{p+q} \rho(y_1 - q)\rho(y_2 - y_1) \cdots \rho(y_n - y_{n-1})\, dy \cdots dy_n. \end{aligned}$$

Consider the symmetric integral equation

$$(4) \qquad \int_0^{p+q} \rho(s - t) f(t)\, dt = \lambda f(s)$$

[6] *Ann. d. Phys.* 53 (1917) p. 177.

and note that if $K_n(s, t)$ denotes the n-th iterated kernel of this integral equation, the right side of (3) is equal to

$$\int_0^{p+q} K_n(q, t)\, dt.$$

Thus

$$\text{Prob. } \{N > n\} = \int_0^{p+q} K_n(q, t)\, dt.$$

From the general theory of integral equations we know that

$$K_n(s, t) = \sum_{j=1}^{\infty} \lambda_j^n f_j(s) f_j(t), \qquad (n \geq 2),$$

where $\lambda_1, \lambda_2, \cdots$ are eigenvalues and $f_1(t), f_2(t), \cdots$ normalized eigenfunctions of the integral equation (4).

Since ρ was assumed to be continuous it follows that the eigenfunctions are continuous and

$$\text{Prob. } \{N > n\} = \sum_{j=1}^{\infty} \lambda_j^n f_j(q) \int_0^{p+q} f_j(t)\, dt.$$

This formula is very general and provides, in a sense, a complete solution of the problem in the continuous and symmetric case. Unfortunately the usefulness of this formula is limited by the difficulties encountered in solving integral equations of the type (4).

In fact, the integral equation

$$\frac{1}{\sqrt{2\pi}} \int_0^a e^{-(s-t)^2/2} f(t)\, dt = \lambda f(s),$$

to which one is led by considering the normally distributed X's, appears to be very difficult to solve.

4. A particular case. If we assume

$$\rho(x) = \frac{1}{2} e^{-|x|}$$

we are led to the integral equation

$$(5) \qquad \int_0^{p+q} e^{-|s-t|} f(t)\, dt = 2\lambda f(s),^{7}$$

which is quite easy to solve.

In fact, rewriting (5) in the form

$$(6) \qquad e^{-s} \int_0^{p+q} e^{t} f(t)\, dt + e^{s} \int_0^{p+q} e^{-t} f(t)\, dt = 2\lambda f(s)$$

[7] I have recently encountered the integral equation (5) in solving an entirely different problem. A complete discussion can be found in a restricted N.D.R.C. Report 14-305.

and differentiating twice with respect to s we obtain the differential equation

$$f''(s) + \left(\frac{1}{\lambda} - 1\right) f(s) = 0.$$

Substituting the general solution of this equation in (6) we find in an entirely elementary fashion that

$$\lambda_j = \frac{1}{1 + y_j^2},$$

$$f_j(t) = \frac{\sin y_j t + y_j \cos y_j t}{\sqrt{1 + \frac{1}{2}(p + q)(1 + y_j^2)}},$$

where y_j is the jth (positive) root of the transcendental equation

$$\tan (p + q)y = -\frac{2y}{1 - y^2}. \tag{7}$$

We have

$$\int_0^{p+q} (\sin y_j t + y_j \cos y_j t)\, dt = \frac{1}{y_j} \{1 - \cos (p + q)y_j + y_j \sin (p + q)y_j\}$$

and it is easily seen that (7) implies

$$1 - \cos (p + q)y_j + y_j \sin (p + q)y_j = \begin{cases} 0 & \text{if } \cos (p + q)y_j = \dfrac{1 - y_j^2}{1 + y_j^2}, \\ 2 & \text{if } \cos (p + q)y_j = -\dfrac{1 - y_j^2}{1 + y_j^2}. \end{cases}$$

Finally,

$$\text{Prob. } \{N > n\} = 2 \sum_{j=1}^{\infty}{}' \frac{1}{(1 + y_j^2)^n} \frac{\sin y_j q + y_j \cos y_j q}{y_j \{1 + \frac{1}{2}(p + q)(1 + y_j^2)\}},$$

where the dash on the summation sign indicates that only those j's are taken under account for which

$$\cos (p + q)y_j = -\frac{1 - y_j^2}{1 + y_j^2}.$$

We omit here the discussion of various limiting cases inasmuch as our main purpose was to obtain exact formulas.

There are indications that some of the limiting cases are related to singular integral equations with continuous spectra. We may return to this subject at a later date.

ON THE CLASSIFICATION OF OBSERVATION DATA INTO DISTINCT GROUPS

By R. v. Mises

Harvard University

Introduction. In scholastic examinations as well as in the examination of industrial products the following probability problem arises. The individuals of a certain population are successively subjected to trials each of which leads to a definite score x (one real number or a group of m real numbers). Each individual is supposed to belong to one of n classes. These classes are characterized by n probability densities $p_1(x)$, $p_2(x)$, $\cdots$ $p_n(x)$. One has to decide on the basis of the observed value x to which class the respective individual belongs and one wishes to make this decision with the smallest possible risk of failure.

For example, let us consider an examination where the three grades A, B, C are attributed on the basis of a simple score x (case $m = 1$, $n = 3$). It may be assumed that an individual of the class A has a mean expected value of x equal to $\vartheta_1 = 75$ and a normal distribution with the standard deviation $\sigma_1 = 4/\sqrt{2}$. The analogous values for the classes B and C may be $\vartheta_2 = 50$, $\sigma_2 = 8/\sqrt{2}$ and $\vartheta_3 = 25$, $\sigma_3 = 12/\sqrt{2}$. In this case, the solution developed in the present paper allows the conclusion that the best way of grading would be to attribute the grade A to scores x beyond 70.0, the grade C to scores below 40.0 and B to the rest. The corresponding error risk will be 3.9% or the success rate 0.961.

There exists, of course, one case where the solution is trivial. If the probability densities $p_\nu(x)$ are limited to n non-overlapping regions R_ν (with $p_\nu = 0$ at points outside R_ν) an obvious decision can be made without any risk of failure. An assumption of this kind underlies the usual procedure of grading. If, in the foregoing example, an individual of class A is supposed to have at any rate a score beyond 60 and a class C individual less than 40, it is obvious how the grades should be attributed without incurring any risk. It seems, however, that in many problems the assumption of normal distributions or some other kind of overlapping distributions is more appropriate. Then, the probability problem has to be solved.

The solution submitted in the present paper is derived from the simplest principles of calculus of probability without any arbitrary assumption or hypothesis. If n equals 2, the problem can also be considered as a problem of testing a simple statistical hypothesis with a two-valued parameter.[1] It has been shown in an earlier paper[2] that under this restriction success rates higher than 50% are obtainable.

---

[1] See A. Wald, *Annals of Math. Stat.*, Vol. 15 (1944), p. 145. Here, both $p_1(x)$ and $p_2(x)$ are supposed to be normal distributions with the same covariance matrix. The problem treated by Wald is different from the one considered in the present paper since in Wald's paper the parameters of the two multivariate normal distributions are assumed to be unknown.

[2] R. v. Mises, *Annals of Math. Stat.*, Vol. 14 (1943), p. 238.

1. Statement of the problem. For each of n classes of individuals a probability density $p_\nu(x)$, $\nu = 1, 2, \cdots n$, is given. We subdivide the m-dimensional x-space into n regions $R_1, R_2, \cdots R_n$ and assign the region R_ν to the νth class. The probability, for an individual of this class, to have its x-value falling in R_ν is

$$(1) \qquad P_\nu = \int_{(R_\nu)} p_\nu(x)\, dX, \qquad \nu = 1, 2, \cdots n$$

where dX denotes the element of the x-space ($dX = dx$ in the case $m = 1$). In the N first trials of the indefinite sequence of trials, N_ν individuals that belong to the νth class will be tested. Out of these only those individuals whose x-value falls in R_ν will be ascribed to the νth class. Their number according to the definition of probability, equals $N_\nu(P_\nu + \epsilon_\nu)$ where ϵ_ν tends towards zero as N_ν goes to infinity. The total number of correct decisions during the N first trials is therefore

$$(2) \qquad N_1(P_1 + \epsilon_1) + N_2(P_2 + \epsilon_2) \cdots N_n(P_n + \epsilon_n)$$

and the relative number is

$$(2') \qquad \frac{N_1}{N}(P_1 + \epsilon_1) + \frac{N_2}{N}(P_2 + \epsilon_2) + \cdots \frac{N_n}{N}(P_n + \epsilon_n).$$

If N increases indefinitely a part of the N_ν must become infinite. For these classes, ϵ_ν converges toward zero. For the other classes N_ν/N diminishes to zero. Thus, the relative number of right decisions converges towards

$$(3) \qquad \frac{1}{N}(N_1P_1 + N_2P_2 + \cdots N_nP_n).$$

The N_ν are unknown. Every one of these unknowns can take each value from zero to N. If P_μ is the smallest P_ν, the most unfavorable case, where the expression (3) has its smallest value, will occur with $N_\mu = N$, all other N_ν being zero. This value is obviously P_μ. Thus it is seen that the frequency of correct assignments is at least equal to the smallest P_ν which may be written as $P_{\min}$. The greatest risk of making a false decision is $1 - P_{\min}$.

Now the problem to be solved in the present paper can be stated as follows: *For n given densities $p_\nu(x)$, find the subdivision of the x-space into n regions R_ν that gives to the smallest of the expressions P_ν defined in* (1) *its possibly greatest value.*

This problem has the type of a continuous variation problem with the integrals in question bounded within the limits zero to one. We may, therefore, assume that under reasonable restrictions for $p_\nu(x)$ a solution exists. Uniqueness of the solution cannot be expected in general. It seems very difficult to establish the conditions for unicity in other than the most simple cases. Existence of more than one solution would mean that each of them is an optimum with respect to infinitesimal modifications of the boundaries.

2. General solution. A simple problem of variation is considered as solved in principle when the nature of the extremals is known. In our case of a so-called minimax problem, where the minimum of n quantities is maximized, an additional relation between the n integrals is required. Both can easily be found in the actual case.

Let us first consider a partition of the x-space into n regions with not all P_ν being equal. The smallest P_ν will be called $P_{\min}$ and the smallest but one P^*. Among the k regions for which $P_\nu = P_{\min}$ there will be at least one, say, R_α that has a common border with a region R_β whose P-value is greater, so that $P_\beta \geqq P^*$. Now modify the boundary between R_α and R_β in such a way that the space covered by R_α is increased and that of R_β decreased. According to (1) the new values of P_α and P_β will be

$$P'_\alpha = P_\alpha + \Delta, \qquad P'_\beta = P_\beta - \Delta' \tag{4}$$

with both Δ and Δ' positive. The two quantities Δ and Δ' are not independent of one another, but they can be chosen both smaller than any given positive number ϵ. Therefore, the condition

$$P'_\alpha = P_\alpha + \Delta < P_\beta - \Delta' = P'_\beta \tag{5}$$

can be fulfilled. All other P_ν-values remain unchanged.

In the case $k = 1$, that is, if only one region R_ν had originally the minimum P-value, the modified system has a greater minimum P, which equals either $P_\alpha + \Delta$ or P^*. If $k > 1$ the new system has the same minimum P as the original one, but its k-value is diminished by one. If we repeat the same procedure $(k - 1)$ times we obtain a system of regions with one single P_ν having the minimum P-value and the next step leads to a partition of the x-space into n regions with a smallest P-value that is greater than the original $P_{\min}$. Thus it is seen that no partition with unequal P_ν-values can solve our problem.

Secondly, if $m > 1$, consider a system of n regions with $P = P_1 = P_2 = \cdots = P_n$. Take two points, x and y, on the border of any two neighboring regions R_ν and R_μ. An infinitesimal variation of the boundary would consist of adding to R_ν in the neighborhood of the point x a space element δS subtracting it from R_μ and, at the same time, adding to R_μ in the vicinity of y an element $\delta S'$ subtracting it from R_ν. Then, according to (1), the new values of P_ν and P_μ will be

$$\begin{aligned} P'_\nu &= P + p_\nu(x)\delta S - p_\nu(y)\delta S' \\ P'_\mu &= P - p_\mu(x)\delta S + p_\mu(y)\delta S'. \end{aligned} \tag{6}$$

Introducing $\Delta_\nu = P'_\nu - P$ and $\Delta_\mu = P'_\mu - P$, these equations solved for δS and $\delta S'$ give

$$\delta S = \frac{p_\nu(y)\Delta_\mu + p_\mu(y)\Delta_\nu}{D}, \qquad \delta S' = \frac{p_\nu(x)\Delta_\mu + p_\mu(x)\Delta_\nu}{D} \tag{7}$$

where

$$D = p_\nu(x)p_\mu(y) - p_\mu(x)p_\nu(y). \tag{7'}$$

If the determinant D is positive, we find two positive quantities δS and $\delta S'$ for any pair of positive Δ_μ and Δ_ν. If D is negative the same is true when x and y are interchanged. In both cases, that is, with $D \neq 0$, the original partition is replaced by a new system of regions in which only two regions, R_ν and R_μ, have increased P-values, while (if $n > 2$) still $P_{\min} = P$. If to this system the procedure as described in the foregoing is applied, a final partition with a greater minimum value of P can be derived. The conclusion is that no solution of our problem can include a boundary on which the determinant D is different from zero for any two points x and y. On the other hand, it is seen that $D = 0$ means that the ratio $p_\nu(x):p_\mu(x)$ has a constant value along the border. Thus the result is reached:

The partition of the x-space that solves our problem is characterized by two properties: (1) *for all n regions R_ν the value of P_ν is the same;* (2) *along the border between R_ν and R_μ the ratio $p_\nu(x)/p_\mu(x)$ is constant.*

In the one-dimensional case ($m = 1$) only the first of these two statements is relevant. In any case, the success rate, that is, the guaranteed ratio of correct decisions, equals the common value of all P_ν.

3. Illustrations. (a) One-dimensional case. Upon introducing the cumulative distribution functions

$$F_\nu(x) = \int_{-\infty}^{x} p_\nu(z)dz \tag{8}$$

the conditions $P_1 = P_2 = \cdots P_n$ take the form

$$F_1(x_1) = F_2(x_2) - F_2(x_1) = \cdots = F_{n-1}(x_{n-1}) - F_{n-1}(x_{n-2}) = 1 - F_n(x_{n-1}) \tag{9}$$

where $x_1, x_2, \cdots x_{n-1}$ determine the n intervals on the both-sides infinite x-axis. If all density functions have the same form except for an affine transformation, one has

$$F_\nu(x) = F[h_\nu(x - \vartheta_\nu)], \qquad \nu = 1, 2, \cdots n \tag{10}$$

Let us assume, for instance, that scores between 0 and 100 are attributed to three types of individuals. The first type may have an even chance to obtain a score between 0 and 50, the second between 40 and 80 and the third between 70 and 100. Here

$$F_\nu(x) = \tfrac{1}{2} + (x - \vartheta_\nu)p_\nu, \qquad |x - \vartheta_\nu| \leqq \frac{1}{2p_\nu} \tag{11}$$

with $\vartheta_\nu = 25, 60, 85$ and $p_\nu = \frac{1}{50}, \frac{1}{40}, \frac{1}{30}$. The conditions (9) supply

$$\frac{1}{2} + \frac{x_1 - 25}{50} = \frac{1}{40}(x_2 - x_1) = \frac{1}{2} - \frac{x_2 - 85}{30} \tag{12}$$

and this, solved for x_1, x_2 gives $x_1 = 41\,\frac{2}{3}$, $x_2 = 75$ while the three expressions (12) take the value 0.833. Therefore, in attributing all scores below $41\frac{2}{3}$ to the first class and all scores beyond 75 to the third one is safe to make under no circumstances more than $\frac{1}{6}$ incorrect decisions in the long run.

In the example quoted in the introduction one has

$$p_\nu(x) = \frac{1}{\sigma_\nu\sqrt{2\pi}}\, e^{(x-\vartheta_\nu)^2/2\sigma_\nu^2} \tag{13}$$

with $\vartheta_\nu = 75, 50, 25$ and $\sigma_\nu^2 = 8, 32, 72$. If $\Phi(x)$ denotes the integral

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2}\, dz$$

the conditions (9) become

$$1 + \Phi\left(\frac{x_1 - 25}{12}\right) = \Phi\left(\frac{x_2 - 50}{8}\right) - \Phi\left(\frac{x_1 - 50}{8}\right) = 1 - \Phi\left(\frac{x_2 - 75}{4}\right). \tag{14}$$

The first and last expression equated lead to $x_1 + 3x_2 = 250$. The complete solution can be found with the help of tables for Φ. It is $x_1 = 29.9920$, $x_2 = 70.0027$ with the common value twice 0.961 for the three expressions (14). Hence the result as quoted in the introduction.

Let us now take up the case of six normal distributions with equidistant mean values $\vartheta = \pm a, \pm 3a, \pm 5a$ and one and the same variance σ^2. Then, because of symmetry, two equations only have to be fulfilled:

$$1 + \Phi\left(\frac{x_1 + 5a}{\sigma\sqrt{2}}\right) = \Phi\left(\frac{x_2 + 3a}{\sigma\sqrt{2}}\right) - \Phi\left(\frac{x_1 + 3a}{\sigma\sqrt{2}}\right) = \Phi\left(\frac{a}{\sigma\sqrt{2}}\right) - \Phi\left(\frac{x_2 + a}{\sigma\sqrt{2}}\right)$$

For $\sigma^2/a^2 = 0.32$, the numerical solution gives

$$x_1 = -4.160a, \qquad x_2 = -2.062a.$$

The success rate, i.e. half the common value of the above expressions is 0.931. The six intervals extend from $-\infty$ to x_1, from x_1 to x_2, from x_2 to 0, from 0 to $-x_2$, from $-x_2$ to $-x_1$, and from $-x_1$ to ∞.

(b) *Case of more than one dimension.* Let us assume that two classes A and B have uniform distributions extending over volumes $V_1 = 1/p_1$ and $V_2 = 1/p_2$ respectively. If the two regions have a common part of volume V each surface within the common space fulfills the condition $p_1/p_2 = $ constant. Thus, the two regions R_1 and R_2 are not uniquely determined but subject to one condition only which determines the optimum success rate. If κV is cut out from V_1 and $(1 - \kappa)V$ from V_2, the relation must be fulfilled:

$$1 - p_1 V\kappa = 1 - p_2 V(1 - \kappa), \qquad i.e.\ \kappa = \frac{p_2 V}{p_1 + p_2}$$

and the success rate is

$$S = 1 - p_1 V\kappa = 1 - \frac{p_1 p_2 V}{p_1 + p_2} = 1 - p_2 V(1 - \kappa).$$

If three classes A, B, and C are considered with the densities $p_1 = 1/V_1$, $p_2 = 1/V_2$, $p_3 = 1/V_3$ and the first two regions have a space of volume V in common, the latter two a space of volume V', the conditions are

$$1 - p_1 V(1 - \kappa) = 1 - p_2(\kappa V + \lambda V') = 1 - p_3(1 - \lambda)V'$$

which supply

$$\kappa = 1 - \frac{p_2 + p_3}{p_1 p_2 + p_2 p_3 + p_3 p_1} \frac{V + V'}{V},$$

$$\lambda = 1 - \frac{p_1 p_2 p_3}{p_1 p_2 + p_2 p_3 + p_3 p_1} \frac{V + V'}{V'}$$

and the success rate has the value

$$S = 1 - (V + V') \frac{p_1 p_2 p_3}{p_1 p_2 + p_2 p_3 + p_3 p_1}.$$

If the p_ν are normal density functions, say

$$p_\nu(x, y) = \frac{\sqrt{D_\nu}}{\pi} e^{-\frac{1}{2}Q_\nu},$$

$$Q_\nu = \alpha_\nu (x - a_\nu)^2 + 2\beta_\nu (x - a_\nu)(y - b_\nu) + \gamma_\nu (y - b_\nu)^2$$

and D_ν the corresponding determinants, the curves separating the regions R_ν are the conics

$$Q_\nu - Q_\mu = \text{const.}$$

where the constants are determined by the conditions that all P_ν must be equal. If the α, β, γ have the same values for every ν, the borders consist of straight lines. In this case one can reduce the expressions for p_ν, by an affine transformation, to

$$p_\nu(x, y) = \frac{1}{\pi} e^{-(x-a_\nu)^2-(y-b_\nu)^2}.$$

In the transformed plane the borderline between the regions R_ν and R_μ is perpendicular to the straight line that connects the points $A_\nu(a_\nu, b_\nu)$ and $A_\mu(a_\mu, b_\mu)$. If all points A_ν lie on the same straight line (in particular, if $n = 2$) the whole problem is practically identical with the one-dimensional ($m = 1$). In the case $n = 3$, in general, the three regions are confined by three lines perpendicular to A_1A_2, A_2A_3, A_3A_1 passing through a point C whose coordinates are determined by the equations $P_1 = P_2 = P_3$. If r_ν denotes the distance $A_\nu C$ and φ_ν, ϑ_ν are the angles, $A_\nu C$ forms with the adjacent sides of the triangle $A_1A_2A_3$ one has to use the function

$$F(r, \varphi_1) = \frac{1}{2\sqrt{\pi}} \int_0^\infty \phi(r - z \tan \varphi) e^{-z^2} \, dz.$$

Then the two conditions for C read

$$F(r_1, \varphi_1) + F(r_1, \vartheta_1) = F(r_2, \varphi_2) + F(r_2, \vartheta_2) = F(r_3, \varphi_3) + F(r_3, \vartheta_3)$$

and the success rate equals 0.5 plus the common value of these three expressions.

ON AN EXTENSION OF THE CONCEPT OF MOMENT WITH APPLICATIONS TO MEASURES OF VARIABILITY, GENERAL SIMILARITY, AND OVERLAPPING[1]

MILTON DA SILVA RODRIGUES
State University of São Paulo

1. Introduction. Given a frequency distribution D: $[X_i, F_i]$ $(i = 1, 2, 3, \cdots, n)$, we shall call the expression

$$M_r(D, X_j) = \sum_{i=1}^{n} (X_i - X_j)^r F_i$$

the rth total moment of D about the origin X_j. We shall consider the weighted sum

$$\mathfrak{M}_r = \Sigma_j W_j M_r(D, X_j)$$

where W_j denotes the weight corresponding to the particular origin X_j, and the summation is over a field ϕ. In particular, if ϕ is the set of all values assumed in D by the variate X_i, and if $W_j = F_j$, we shall call the quantity the rth *complete total moment* of D. If, on the contrary, W_j is the frequency F'_j of the value X'_j in a second frequency distribution D': $[X'_j, F'_j]$ and ϕ' is the set of all values assumed by the variate X'_j in D', $\mathfrak{M}_r$ will be called the rth *aggregate moment* of D and D'. A modification of this procedure leads to what we shall call the *moment of transvariation* of D and D'.

The consideration of complete moments draws attention to certain previously known measures of variability which are independent of the origin selected, and also provides simple methods of computation which are useful for data given in the form of a frequency distribution. The investigation of aggregate moments and moments of transvariation gives rise to certain measures of general similarity between two distributions, as well as measures of the amount of overlapping.

2. Sliding and complete moments of a frequency distribution.

2.1. We shall give the name *sliding total moments* of order r to the successive values, for particular values of j, of the expression

$$M_r(X_j) = F_j \sum_{i=1}^{n} [(X_i - X_j)^r F_i]. \tag{2.11}$$

[1] The Portuguese original of this paper was written in Brazil, in August 1943. Its translation into English was entirely revised by Dr. T. Greville, Bureau of the Census, who proposed also many simplifications in the derivation of formulae. For his painstaking labor and interest I wish to express my very sincere appreciation. I also wish to thank Dr. W. Edwards Deming for reading the manuscript and making several valuable suggestions.

The expression for the complete total moment, written out in full, is

$$\mathfrak{M}_r = \sum_{j=1}^{n} M_r(X_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} [(X_i - X_j)^r F_i F_j]. \tag{2.12}$$

It is readily seen that the complete moment is independent of the choice of origin.

2.2. If $r = 0$, we have

$$M_0(X_j) = F_j \sum_{i=1}^{n} F_i .$$

The complete total moment of order zero will therefore be

$$\mathfrak{M}_0 = \sum_{j=1}^{n} F_j \sum_{i=1}^{n} F_i = M_0^2 \tag{2.21}$$

where M_0 stands for the total moment of order zero about the origin of the X', that is,

$$M_0 = N\nu_0'.$$

2.3. If $r = 1$, we shall have

$$M_1(X_j) = F_j \sum_{i=1}^{n} [(X_i - X_j) F_i].$$

Using M_1 to denote the total moment of order one about the origin of the X, we obtain

$$M_1(X_j) = F_j \sum_{1}^{n} X_i F_i - X_j F_j \sum_{1}^{n} F_i = F_j M_1 - X_j F_j M_0 .$$

Making j vary from 1 to n and summing, we have

$$\begin{aligned} \mathfrak{M}_1 &= \sum_{j=1}^{n} F_j M_1 - \sum_{j=1}^{n} X_j F_j M_0 \\ &= M_0 M_1 - M_1 M_0 = 0. \end{aligned} \tag{2.31}$$

This result is due to the fact that we took the deviations $X_i - X_j$ with their proper signs. We may, however, calculate the value which the complete moment of first order would have if using absolute values. Thus, the sliding total moment thus modified becomes

$$| M_1(X_j) | = F_j \left[\sum_{i=1}^{j-1} (X_j - X_i)F_i + \sum_{i=j}^{n} (X_i - X_j)F_i \right]$$

which may be put in the form

$$| M_1(X_j) | = F_j X_j \left[\sum_{i=1}^{j-1} F_i - \sum_{i=j}^{n} F_i \right] - F_j \left[\sum_{i=1}^{j-1} F_i X_i - \sum_{i=j}^{n} F_i X_i \right]. \tag{2.32}$$

Summing with respect to j and employing the substitutions

$$\sum_{i=j}^{n} F_i = M_0 - \sum_{i=1}^{j-1} F_i \tag{2.33}$$
$$\sum_{i=j}^{n} F_i X_i = M_1 - \sum_{i=1}^{j-1} F_i X_i$$

gives for the complete total moment

$$|\mathfrak{M}_1| = 2\sum_{j=1}^{n}\left[F_j X_j \sum_{i=1}^{j-1} F_i\right] - 2\sum_{j=1}^{n}\left[F_j \sum_{i=1}^{j-1} F_i X_i\right]. \tag{2.34}$$

The quotient

$$m_1 = \frac{|\mathfrak{M}_1|}{\mathfrak{M}_0} \tag{2.35}$$

of the complete total moment of order one by the complete total moment of order zero we shall call the *complete unit moment* of order one, or simply the complete moment of order one, when no confusion would result.

The complete unit moment is a measure of variability, identical with that already considered by Andrae and Helmert, respectively in 1869 and in 1876, and which C. Gini, in 1912, called mean difference with repetition.[2]

The numerator of m_1 is easily computed if we observe that the upper limit $j - 1$ of the F_i summation, for example, means that each product $X_j F_j$ must be multiplied by the cumulative frequency corresponding to the class immediately preceding. We only have to shift the cumulative frequencies column by one class in the proper direction; the second term is similarly dealt with.

2.4. The second order sliding total moment is

$$M_2(X_j) = F_j \sum_{i=1}^{n} [(X_i - X_j)^2 F_i] = F_j M_2 - 2F_j X_j M_1 + F_j X_j^2 M_0$$

where M_2 is the total moment of order two. Summing with respect to j gives the complete total moment of order two

$$\mathfrak{M}_2 = \sum_{j=1}^{n} M_2(X_j) = 2(M_2 M_0 - M_1^2). \tag{2.41}$$

The complete unit moment of order two is therefore

$$\begin{aligned} m_2 &= 2\left[\frac{M_2}{M_0} - \left(\frac{M_1}{M_0}\right)^2\right] = 2(\nu_2' - \nu_1'^2) \\ &= 2\sigma^2 \end{aligned} \tag{2.42}$$

---

[2] Apud Czuber, *Wahrscheinlichkeitsrechnung*, Vol. 2, (1932), p. 316. C. Gini, *Variabilità e Mutabilità*, Cagliari, 1912.

where ν' stands for a unit moment about the origin of the X, namely

$$\nu_r' = \frac{\Sigma X^r F}{\Sigma F},$$

m_2 is also a measure of variability, independent of the choice of origin. It is equal to the square of Gauss's "Präzisionsmass", and to the double of Fisher's variance; like m_1 it was defined by Andrae and Helmert, and was called by Gini the mean square difference with repetition.

2.5. If $r = 3$ we have for the sliding moments,

$$\begin{aligned} M_3(X_j) &= F_j \sum_{i=1}^{n} (X_i - X_j)^3 F_i \\ &= F_j M_3 - 3F_j X_j M_2 + 3F_j X_j^2 M_1 - F_j X_j^3 M_0 . \end{aligned}$$

Summation over j gives

$$(2.51) \qquad \mathfrak{M}_3 = \sum_{j=1}^{n} M_3(X_j) = M_0 M_3 - 3M_1 M_2 + 3M_2 M_1 - M_3 M_0 = 0,$$

a result which is easily shown to hold for any complete moment of odd order. We may calculate the value of the complete moment of order three using absolute values of the deviations $X_i - X_j$ by a process similar to that previously described for the calculation of $|\mathfrak{M}_1|$. This gives

$$(2.52) \qquad \begin{aligned} |\mathfrak{M}_3| = 2\Bigg[\sum_{j=1}^{n} F_j X_j^3 \sum_{i=1}^{j-1} F_i - 3 \sum_{j=1}^{n} F_j X_j^2 \sum_{i=1}^{j-1} F_i X_i \\ + 3 \sum_{j=1}^{n} F_j X_j \sum_{i=1}^{j-1} F_i X_i^2 - \sum_{j=1}^{n} F_j \sum_{i=1}^{j-1} F_i X_i^3\Bigg]. \end{aligned}$$

2.6. The sliding moments of order four are

$$M_4(X_j) = F_j M_4 - 4F_j X_j M_3 + 6F_j X_j^2 M_2 - 4F_j X_j^3 M_1 + F_j X_j^4 M_0 .$$

Summing with respect to j and simplifying, we have

$$(2.61) \qquad \begin{aligned} \mathfrak{M}_4 &= M_0 M_4 - 4M_1 M_3 + 6M_2^2 - 4M_3 M_1 + M_4 M_0 \\ &= 2(M_0 M_4 - 4M_1 M_3 + 3M_2^2). \end{aligned}$$

Dividing both sides by $\mathfrak{M}_0$ in order to obtain the complete moment on a unit basis, we have

$$m_4 = 2\left[\frac{M_4}{M_0} - 4\frac{M_1}{M_0}\frac{M_3}{M_0} + 3\left(\frac{M_2}{M_0}\right)^2\right] = 2\,(\nu_4' - 4\nu_1'\nu_3' + 3\nu_2'^2).$$

But, if ν indicates a moment about the mean

$$\nu_4 = \nu_4' - 4\nu_1'\nu_3' + 6\nu_1'^2\nu_2' - 3\nu_1'^4.$$

By substitution, therefore

$$(2.62)\qquad \begin{aligned} m_4 &= 2(\nu_4 + 3\nu_2'^2 - 6\nu_1'^2\nu_2' + 3\nu_1'^4) \\ &= 2[\nu_4 + 3(\nu_2' - \nu_1'^2)^2] \\ &= 2(\nu_4 + 3\nu_2^2). \end{aligned}$$

This complete moment gives rise to a measure of kurtosis independent of the choice of origin

$$k = \frac{m_4}{m_2^2} = \frac{\nu_4}{2\nu_2^2} + \frac{3}{2}.$$

In case of mesokurtosis this reduces to 3, since for the normal curve $\nu^4/\nu_2^2 = 3$; leptokurtosis and platikurtosis occur for the same ranges as in the case of Pearson's measure β_2.

3. Aggregate moments of two frequency distributions.

3.1. Given two frequency distributions, D:$[X_i, F_i](i = 1, 2, 3, \cdots, n)$ and D': $[X_j', F_j'](j = 1, 2, 3, \cdots, p)$ and a fixed point X_j' belonging to the second distribution, we shall call the expression

$$(3.11)\qquad M_r(D, X_j') = F_j' \sum_{i=1}^{n} (X_i - X_j')^r F_i$$

the rth *aggregate sliding total moment* of the first distribution about the element X_j' of the second. Summation over j gives

$$(3.12)\qquad {}^c\mathfrak{M}_r = \sum_{j=1}^{p} \sum_{i=1}^{n} F_j'(X_i - X_j')^r F_i.$$

We shall call ${}^c\mathfrak{M}_r$ the *aggregate complete total moment* or, simply, the aggregate total moment of D about D'. It is clear that this is a symmetric function of the two distributions, except for a change of sign in the case of odd moments.

3.2. If $r = 0$, we have

$$(3.21)\qquad M_0(D, X_j') = F_j' \sum_{i=1}^{n} F_i$$

$$(3.22)\qquad {}^c\mathfrak{M}_0 = \sum_{j=1}^{p} F_j' \sum_{i=1}^{n} F_i = M_0 M_0'.$$

3.3. If $r = 1$, we have

$$(3.31)\qquad M_1(D, X_j') = F_j' M_1 - F_j' X_j' M_0$$

$$(3.32)\qquad {}^c\mathfrak{M}_1 = M_1 M_0' - M_0 M_1'.$$

We shall call the quotient

$$(3.33)\qquad {}^c m_1 = \frac{{}^c\mathfrak{M}_1}{{}^c\mathfrak{M}_0}$$

the aggregate *unit* moment of order r (or the aggregate moment coefficient), or simply the aggregate moment of order r whenever the simpler name will not cause confusion.

It is obvious that the aggregate moments are measures of general similarity, as to form and position, between D and D'. This similarity will be an identity in case the two distributions coincide perfectly; on the other hand, it is clear that there is no limit to the degree of non-similarity which may be encountered. We shall take unity to represent the maximum and zero the minimum of similarity, and thus define a provisional similarity index

$$S = \frac{m_1 m_1'}{{}^c m_1^2}. \tag{3.34}$$

But

$${}^c m_1 = \frac{M_1 M_0' - M_0 M_1'}{M_0 M_0'} = A - A'$$

where A and A' stand for the arithmetic means of D and D', respectively. Now it will be seen that if $A = A'$, $S = \infty$. This result is due to the fact that in the calculation of m_1 and m_1' we took the absolute values of the deviations $X_i - X_j'$, while in the calculation of ${}^c m_1$ we retained the algebraic signs. In order to make the two terms of the fraction in (3.34) comparable, we can either: 1) calculate ${}^c m_1$ also using absolute values; or 2) take only the positive or only the negative part of both numerator and denominator of S. In any case, $A = A'$ is a necessary condition for the maximum of S.

3.4. We shall employ the first method suggested above, although we shall return to the second in the third part of the paper. As long as D and D' do not overlap, all the $X_i - X_j'$ deviations have the same sign and this is the same as that of the difference $A - A'$. If, however, there is some overlapping this will not be the case, some deviations having different signs from that of $A - A'$. This brings us to Gini's concept of "transvariation". He applies this term to any deviation $X_i - X_j'$ which does not have the same sign as $\bar{X} - \bar{X}'$, these symbols denoting averages of any previously specified type; and he calls the magnitude of the deviation its "intensity".

In computing the complete moment of the first order using absolute values, in order to simplify the algebra we shall assume the same origin for X and X' and therefore drop the stroke from the X, but not of course from the F. If certain values of X occur in one distribution and not in the other, we can merely consider the frequency as zero in the second distribution. In this way the two distributions can be regarded as extending over the same total range. If X_1 and X_m denote the extreme values, the sliding total moment is

$$\begin{aligned} |M_1(D, X_j)| &= F_j'\left[\sum_{i=1}^{j-1}(X_j - X_i)F_i + \sum_{i=j}^{m}(X_i - X_j)F_i\right] \\ &= F_j' X_j\left(\sum_{i=1}^{j-1} F_i - \sum_{i=j}^{m} F_1\right) - F_j'\left(\sum_{i=1}^{j-1} F_i X_i - \sum_{i=j}^{m} F_i X_i\right). \end{aligned}$$

Summing with respect to j and at the same time employing the substitutions (2.33) or their transposed form, we obtain the following alternative expressions for the complete aggregate moment:

$$(3.41)\quad |{}^c\mathfrak{M}_1| = M_1M_0' - M_0M_1' + 2\sum_{j=1}^{m}\left[F_j'X_j\sum_{i=1}^{j-1}F_i\right] - 2\sum_{j=1}^{m}\left[F_j'\sum_{i=1}^{j-1}F_iX_i\right]$$

$$(3.42)\quad |{}^c\mathfrak{M}_1| = M_0M_1' - M_1M_0' - 2\sum_{j=1}^{m}\left[F_j'X_j\sum_{i=j}^{m}F_i\right] + 2\sum_{j=1}^{m}\left[F_j'\sum_{i=j}^{m}F_iX_i\right].$$

Note the similarity of the first of these forms to formula (2.34) which is in fact a particular case of formula (3.41). Alternatively, we may obtain from formula (3.42) the particular case

$$(2.34a)\qquad |\mathfrak{M}_1| = 2\sum_{j=1}^{n}\left(F_j\sum_{i=j}^{n}F_iX_i\right) - 2\sum_{j=1}^{n}\left(F_jX_j\sum_{i=j}^{n}F_i\right)$$

which is equivalent to (2.34).

If the two distributions do not overlap, $|{}^c\mathfrak{M}_1|$ does not differ numerically from ${}^c\mathfrak{M}_1$. Let us consider the case in which there is actual overlapping, the range of non-zero frequencies extending from X_1 to X_{n+p} for D and from X_{n+1} to X_m for D'. Then formula (3.42) becomes, upon merely dropping all vanishing terms

$$(3.43)\quad \begin{aligned} |{}^c\mathfrak{M}_1| &= M_0M_1' - M_1M_0' \\ &\quad - 2\sum_{j=n+1}^{n+p}\left[F_j'X_j\sum_{i=n+1}^{j-1}F_i\right] + 2\sum_{j=n+1}^{n+p}\left[F_j'\sum_{i=j}^{n+p}F_iX_i\right]. \end{aligned}$$

On the other hand, formula (3.41) reduces, under the same circumstances, to a much less simple expression, which upon making the substitutions (2.33) and simplifying reduces to

$$(3.44)\quad \begin{aligned} |{}^c\mathfrak{M}_1| &= M_0M_1' - M_1M_0' + 2\sum_{j=n+1}^{n+p}\left[F_j'X_j\sum_{i=n+1}^{j-1}F_i\right] \\ &\quad - 2\sum_{j=n+1}^{n+p}\left[F_j'\sum_{i=j}^{n+p}F_iX_i\right] \\ &\quad - 2\sum_{j=n+1}^{n+p}F_j'X_j\sum_{i=n+1}^{n+p}F_i + 2\sum_{j=n+1}^{n+p}F_j'\sum_{i=n+1}^{n+p}F_iX_i. \end{aligned}$$

This result may be arrived at somewhat more easily by merely making the substitutions (2.33) directly in formula (3.43). It may be noted that formula (3.44) at once reduces to the form (2.34) if the two distributions are identical, since the additional terms all cancel. It is, however a less satisfactory result than formula (3.43) because of the larger number of terms it contains. In order to obtain a formula which resembles (2.34) more closely, we may reverse the

order of summation in formula (3.43). Observing that the terms for $j = i$ collectively vanish, we see that

$$|{}^c\mathfrak{M}_1| = M_0 M_1' - M_1 M_0' - 2 \sum_{i=n+1}^{n+p} \left[F_i \sum_{j=n+1}^{i-1} F_j' X_j \right] + 2 \sum_{i=n+1}^{n+p} \left[F_i X_i \sum_{j=n+1}^{i-1} F_j' \right]. \tag{3.45}$$

It will be seen that the simple method of numerical computation described in section 2.8 is immediately applicable to all the formulas (3.41) to (3.45). Dividing any of these expressions by ${}^c\mathfrak{M}_0$ gives $|{}^c m_1|$. For example, if formula (3.43) is used, we have

$$|{}^c m_1| = A' - A - \frac{2}{M_0 M_0'} \left\{ \sum_{j=n+1}^{n+p} \left[F_j' X_j \sum_{i=j}^{n+p} F_i \right] - \sum_{i=j}^{n+p} \left[F_j' \sum_{i=j}^{n+p} F_i X_i \right] \right\}. \tag{3.46}$$

Substituting this value in equation (3.34), we have

$$S_1 = \frac{m_1 m_1'}{|{}^c m_1|^2} \tag{3.47}$$

a quantity which we shall call the "mean coefficient of similarity."

We now observe that S_1 is a general measure of similarity whose magnitude is affected by differences in either form or position. It may, however, be desirable to eliminate the position element, in order to isolate the form aspect. To do this it will suffice to relate the value which $|{}^c m_1|$ would have for $A = A'$, to the product $m_1 m_1'$. This value of $|{}^c m_1|$ is, in fact, its minimum; denoting it by ${}^c\mu_1$ we obtain the index

$$\mathfrak{S}_1 = \frac{m_1 m_1'}{{}^c\mu_1^2} \tag{3.48}$$

which we shall call the mean similarity ratio.

It is clear that all the above mentioned indices measure overlapping as well as similarity. Overlapping between two distributions will be greatest when their similarity is greatest, or when $|{}^c m_1|$ is a minimum. In order to bring out more clearly the overlapping aspect we may follow Gini's procedure of contrasting the actual value of a measure with its maximum value. As already pointed out, if the form of the two distributions is held constant, but their relative position is varied, the degree of overlapping, as measured by the mean similarity ratio, is greatest when the arithmetic means coincide. This method of procedure is embodied in the index

$$\mathfrak{T}_1 = \frac{{}^c\mu_1}{{}^c m_1} \tag{3.49}$$

which we shall call the "intensity of transvariation or overlapping." To calculate ${}^c\mu_1$ we may, for example, merely add the difference $A' - A = c$ to the X

values, in order to move D along the X-axis a distance of c, and then proceed to calculate $|{}^c m_1|$ in the usual manner from the adjusted X values.

3.5. If, in (3.11), $r = 2$, we have

$$M_2(D, X_j) = F'_j \sum_{i=1}^{n} (X_i - X_j)^2 F_i$$
$$= F'_j M_2 - 2X'_j F'_j M_1 + X'^2_j F'_j M_0 .$$

Summing for j then gives

$$(3.51) \qquad {}^c\mathfrak{M}_2 = M'_0 M_2 - 2M'_1 M_1 + M'_2 M_0 .$$

If we define the second aggregate unit moment as

$${}^c m_2 = \frac{{}^c\mathfrak{M}_2}{{}^c\mathfrak{M}_0}$$

then

$$(3.52) \qquad \begin{aligned} {}^c m_2 &= \frac{M_2}{M_0} - 2\frac{M_1}{M_0}\frac{M'_1}{M_0} + \frac{M'_2}{M_0} \\ &= \sigma^2 + \sigma'^2 + (A - A')^2, \end{aligned}$$

where the σ and the A stand for the standard deviations and the arithmetic means of the respective distributions. Now we define the "mean square coefficient of similarity" as the value of

$$(3.53) \qquad \begin{aligned} S_2 &= \frac{m_2 m'_2}{{}^c m_2^2} \\ &= \frac{4\sigma^2 \sigma'^2}{[\sigma^2 + \sigma'^2 + (A - A')^2]^2} . \end{aligned}$$

It is obvious that a minimum value of S_2 requires that $A = A'$ as a necessary condition for the maximum degree of overlapping. Maximum similarity requires, in addition, $\sigma = \sigma'$, in which case $S_2 = 1$.

For a measure of similarity which is independent of difference in position between the two distributions, we define.

$$(3.54) \qquad \mathfrak{S}_2 = \frac{m_2 m'_2}{{}^c \mu_2^2},$$

where ${}^c\mu_2$ is the minimum value of ${}^c m_2$ for all positions of the two distributions, without changing their form. This is obtained by merely taking

$$(3.55) \qquad {}^c\mu_2 = \sigma^2 + \sigma'^2 .$$

For a measure of overlapping we can follow Gini in contrasting the actual

value of ${}^c m_2$ with its minimum ${}^c\mu_2$, since the maximum of overlapping corresponds to the minimum value of ${}^c m_2$. We thus set

$$\mathfrak{T}_2 = \frac{{}^c\mu_2}{{}^c m_2} = \frac{\sigma^2 + \sigma'^2}{\sigma^2 + \sigma'^2 + (A - A')^2} \tag{3.56}$$

a measure which we shall call the "density of overlapping". Its maximum value is unity.

It may be remarked that all the indices proposed in this paragraph are easier to calculate than those of paragraph 3.4. The individual terms are all functions of only one of the two distributions; yet the resulting indices are independent of the origin chosen, and therefore free from any criticism based on doubt as to the representativeness of the arithmetic mean, in cases of marked skewness.

4. Positive and negative moments, and moments of transvariation.

4.1. The aggregate sliding total moment of two frequency distributions D and D' may be expressed in the form

$$M_r(D, X_j') = F_j' \sum_{i=1}^{j-1} (X_i - X_j)^r F_i + F_j' \sum_{i=j+1}^{m} (X_i - X_j)^r F_i \tag{4.11}$$

when both distributions have been artificially extended, if necessary, to cover the same total range, as previously described in section 3.4. We shall characterize the second term in the right member of (4.11) as the positive sliding moment, and the absolute value of the first term as the negative sliding moment. We shall denote these moments by ${}^+M_r(D, X_j)$ and ${}^-M_r(D, X_j)$. The complete moments obtained by summing these separate terms over the range of values of j we shall call the positive and negative aggregate complete moments. Thus the positive complete moment is

$${}^{+c}\mathfrak{M}_r = \sum_{j=1}^{m} \left[F_j' \sum_{i=j+1}^{m} (X_i - X_j)^r F_i \right] \tag{4.12}$$

and the negative complete moment is

$${}^{-c}\mathfrak{M}_r = \sum_{j=1}^{m} \left[F_j' \sum_{i=1}^{j-1} (X_j - X_i)^r F_i \right]. \tag{4.13}$$

That one of these two partial moments which is obtained from differences $X_i - X_j'$ having the opposite sense to that of the difference $\bar{X} - \bar{X}'$ will be called the moment of transvariation of the two distributions and will be denoted by the symbol ${}^T\mathfrak{M}_r$. Here, as in section 3.4, $\bar{X}$ and $\bar{X}'$ denote averages of any previously selected type. For example, if the arithmetic means are the averages selected, and if $A - A'$ is positive, then the negative aggregate moment is the moment of transvariation, and vice-versa.

In the trivial case in which the two distributions are identical, the positive and negative complete moments are equal, and both reduce to merely one half

the aggregate complete moment (computed by the use of absolute values in the case of moments of odd order).

The unit moment of transvariation will be defined as

$$(4.14) \qquad {}^{T}m_r = \frac{{}^{T}\mathfrak{M}_r}{{}^{T}\mathfrak{M}_0}$$

4.2. It is evident that the moments of transvariation can be considered as measures of overlapping. Any such moment equals zero when there is no overlapping and becomes greatest when the two distributions coincide. Taking unity to represent the maximum and zero the minimum of overlapping, we may choose as a general measure of overlapping,

$$(4.21) \qquad T_r = \frac{4\,{}^{T}m_r^2}{|m_r|\,|m_r'|} = \frac{4\,{}^{T}\mathfrak{M}_r^2}{|\mathfrak{M}_r|\,|\mathfrak{M}_r'|}.$$

It will be seen that this quantity always equals zero when there is no overlapping, and equals unity when there is complete overlapping: that is when the two distributions are identical.

5. Need for further developments. All of the measures above described were defined for the case of finite sets of magnitudes, expressed as frequency distributions D and D'. Now these sets of magnitudes may be thought of as samples drawn out of their corresponding universes. The consideration of these universes would lead to more general representations under the form of frequency functions, and the above measures would be expressed as definite integrals rather than summations. This draws attention to the need for tests of significance of the magnitude of all the above measures, especially those of overlapping, in order to allow for sampling fluctuation. Obviously, when the frequency functions are of the asymptotic type some amount of overlapping will always exist.

ON A PROBLEM OF ESTIMATION OCCURRING IN PUBLIC OPINION POLLS

By Henry B. Mann

Ohio State University

To arrive at an estimate of the number of electoral votes that will be cast for a presidential candidate a poll is taken of $\lambda_i N$ interviews in the ith state ($i = 1, \cdots, 48$) where the λ_i are fixed constants > 0 such that $\Sigma\lambda_i = 1$ and the respondent is asked for which candidate he intends to cast his vote. To estimate the number of electoral votes which candidate A will receive, the electoral votes of all the states in which the poll shows a majority for candidate A are added and their sum is used as an estimate for the number of electoral votes which candidate A will receive. In this paper certain properties of this estimate will be discussed. It will be shown that it is a biased but consistent estimate and an upper bound for the bias will be derived. Finally we shall derive that distribution of interviews which minimizes the variance of our estimate.

In all that follows we shall consider the poll as a random or stratified random sample and shall disregard the bias introduced by inaccurate answers. Our results however remain valid as long as the sampling variance is proportional to $\frac{1}{\sqrt{N}}$.

We shall use the following notation:

π_i = proportion of voters in the ith state who intend to vote for candidate A.

$$\epsilon_i = 1 \quad \text{if } \pi_i > \tfrac{1}{2}$$
$$\phantom{\epsilon_i = {}} 0 \quad \text{if } \pi_i < \tfrac{1}{2}$$

w_i = number of electoral votes of the ith state.
p_i, e_i = sample values of π_i and ϵ_i resp.

We shall further exclude the case $\pi_i = \frac{1}{2}$.

The number of electoral votes for candidate A is then given by

$$\sum_{i=1}^{i=48} \epsilon_i w_i = \Gamma.$$

As an estimate of Γ we use the quantity

$$\sum_{i=1}^{i=48} e_i w_i = G. \tag{1}$$

Let ρ_i be the probability that $p_i > \frac{1}{2}$ and hence $e_i = 1$. Let $\lambda_i N = N_i$ be the number of interviews in the ith state. If N_i is not too small then ρ_i is given by

$$\rho_i = \int_{\frac{1}{2}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(x-\pi_i)^2/2\sigma_i^2}\, dx. \tag{2}$$

In this formula $\sigma_i = \sqrt{\frac{\pi_i(1-\pi_i)}{N_i}}$ if the sample is an unstratified random sample and may be somewhat less if the sample is a stratified random sample.[1] For our purposes it is sufficient to assume that σ_i is proportional to $\frac{1}{\sqrt{N_i}}$.

We then have $E(e_i) = \rho_i$ and

$$E(G) = E(\textstyle\sum_{i=1}^{i=48} e_i w_i) = \sum_{i=1}^{i=48} \rho_i w_i. \tag{3}$$

Hence G is a biased estimate of Γ. On the other hand[2] $\underset{N\to\infty}{\text{plim}}\ p_i = \pi_i$ and hence $\underset{N\to\infty}{\text{plim}}\ e_i = \epsilon_i$ and therefore $\underset{N\to\infty}{\text{plim}}\ G = \Gamma$. That is to say G is a consistent estimate of Γ.

According to (3) the bias is given by

$$B(N) = \textstyle\sum_{i=1}^{i=48} \epsilon_i w_i - \sum_{i=1}^{i=48} \rho_i w_i = \sum_{i=1}^{i=48} (\epsilon_i - \rho_i) w_i. \tag{4}$$

We have

$$\epsilon_i - \rho_i = -\frac{1}{\sqrt{2\pi}} \int_{(\frac{1}{2}-\pi_i)/\sigma_i}^{\infty} e^{-\frac{1}{2}x^2}\, dx \quad \text{if} \quad \pi_i < \tfrac{1}{2}$$

$$\epsilon_i - \rho_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\frac{1}{2}-\pi_i)/\sigma_i} e^{-\frac{1}{2}x^2}\, dx \quad \text{if} \quad \pi_i > \tfrac{1}{2}.$$

For a stratified as well as for an unstratified sample σ_i is proportional to $\frac{1}{\sqrt{N_i}}$ and we therefore put

$$\frac{\frac{1}{2} - \pi_i}{\sigma_i} = \begin{cases} \gamma_i \sqrt{N_i} & \text{if} \quad \pi_i < \frac{1}{2} \\ -\gamma_i \sqrt{N_i} & \text{if} \quad \pi_i > \frac{1}{2} \end{cases}. \tag{5}$$

Then we have in both cases

$$|\epsilon_i - \rho_i| = \frac{1}{\sqrt{2\pi}} \int_{\gamma_i\sqrt{N_i}}^{\infty} e^{-\frac{1}{2}x^2}\, dx. \tag{6}$$

We have for $a > 0$

$$\begin{aligned} \int_a^{\infty} e^{-\frac{1}{2}x^2}\, dx &\le h(e^{-\frac{1}{2}a^2} + e^{-\frac{1}{2}(a+h)^2} + e^{-\frac{1}{2}(a+2h)^2} + \cdots) \\ &< e^{-\frac{1}{2}a^2} h(1 + e^{-ah} + e^{-2ah} + \cdots) \\ &= e^{-\frac{1}{2}a^2} \frac{h}{1 - e^{-ah}} \end{aligned}$$

for every value h.

Since $\lim_{h\to 0} \frac{h}{1 - e^{-ah}} = \frac{1}{a}$ we have

$$\int_a^{\infty} e^{-\frac{1}{2}x^2}\, dx \le \frac{e^{-\frac{1}{2}a^2}}{a} \quad \text{for every } a > 0. \tag{7}$$

---

[1] The variance in public opinion polls is somewhat larger than the random sampling variance due to the fact that a cluster sample is used and not a random sample. For the same reason the estimate p_i of π_i may be biased.

[2] For the notation used here see: H. B. MANN AND A. WALD, "On stochastic limit and order relationships". *Annals of Math. Stat.*, (1943), pp. 217–227.

From (6) and (7) we obtain

$$|\epsilon_i - \rho_i| \leq \frac{e^{-\frac{1}{2}\gamma_i^2 N_i}}{\sqrt{2\pi N_i}\,\gamma_i}. \tag{8}$$

From (4) and (8) we have

$$|B(N)| \leq \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{i=48} w_i \frac{e^{-\frac{1}{2}\gamma_i^2 N_i}}{\gamma_i \sqrt{N_i}}. \tag{9}$$

Formula (9) is valid whenever $\pi_i \neq \frac{1}{2}$ and shows that $B(N)$ converges rapidly to 0 for all values $\pi_i \neq \frac{1}{2}$.

To obtain an approximate idea of the magnitude of the bias we may in (4) replace ϵ_i and ρ_i by their sample values e_i and r_i. The quantity $\sum_{i=1}^{i=48} w_i (e_i - r_i)$ can, however, not be regarded as an estimate of $B(N)$.

We now proceed to compute the standard error of G. We may consider the poll as 48 single experiments where the probability of success in the ith experiment is given by ρ_i where

$$\frac{1}{\sqrt{2\pi}} \int_{\gamma_i\sqrt{N_i}}^{\infty} e^{-\frac{1}{2}x^2}\, dx = \begin{cases} \rho_i & \text{if} \quad \pi_i < \frac{1}{2} \\ 1 - \rho_i & \text{if} \quad \pi_i > \frac{1}{2} \end{cases}.$$

Hence the variance of G is given by

$$\sigma^2 = \sum_{i=1}^{i=48} \rho_i (1 - \rho_i) w_i^2. \tag{10}$$

As an estimate of σ^2 we can use the quantity S^2 obtained by replacing ρ_i by its sample value.

We shall consider that distribution of interviews as best which minimizes $E[(G - \Gamma)^2]$.

We have

$$E[(G - \Gamma)^2] = \sigma^2 + B^2(N)$$

We therefore consider the problem of minimizing $\sigma^2 + B^2(N)$ under the restriction $\sum_{i=1}^{i=48} N_i = N$.

We have

$$\frac{\partial \sigma^2}{\partial N_i} = \frac{\partial \sigma^2}{\partial \rho_i} \frac{\partial \rho_i}{\partial N_i} = w_i^2 (1 - 2\rho_i) \frac{\partial \rho_i}{\partial N_i}$$

$$\frac{\partial B^2(N)}{\partial N_i} = 2B(N) \frac{\partial B(N)}{\partial N_i} = -2w_i B(N) \frac{\partial \rho_i}{\partial N_i}$$

$$\frac{\partial \rho_i}{\partial N_i} = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\gamma_i^2 N_i} \frac{\gamma_i}{2\sqrt{N_i}} \quad \text{if} \quad \pi_i < \tfrac{1}{2}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\gamma_i^2 N_i} \frac{\gamma_i}{2\sqrt{N_i}} \quad \text{if} \quad \pi_i > \tfrac{1}{2}.$$

Hence applying the method of Lagrange operators, we obtain

$$(11)\qquad \frac{\partial[\sigma^2 + B^2(N)]}{\partial N_i} = \frac{\partial \rho_i}{\partial N_i} w_i[w_i(1 - 2\rho_i) - 2B(N)] = \lambda, \quad i = 1 \cdots 48.$$

$$\textstyle\sum_{i=1}^{i=48} N_i = N.$$

The parameters γ_i and π_i in equation (11) can be estimated from a previous poll.[3] It is not certain that (11) has always solutions. However if the quantity $\sigma^2 + B^2(N)$ has a minimum for a set of values $N_1, \cdots, N_{48}$ with $N_i \neq 0$ ($i = 1, \cdots, 48$) then (11) must have a solution

One might be induced to try to estimate $\Sigma\, \rho_i w_i$ directly by using $r_i = \frac{1}{\sqrt{2\pi}} \int_{(\frac{1}{2}-p_i)/s_i}^{\infty} e^{-x^2/2}\, dx$ as an estimate of ρ_i. It is easy to see that r_i is a consistent estimate of ϵ_i. It will be shown however that this estimate is more biased than the estimate (1).

Since σ_i differs only very little from its sample estimate s_i we may replace this sample estimate by σ_i. We then have

$$\begin{aligned} E(r_i) &= E\left(\frac{1}{\sqrt{2\pi}\,\sigma_i} \int_{\frac{1}{2}}^{\infty} e^{-(x-p_i)^2/2\sigma_i^2}\, dx\right) \\ &= \frac{1}{2\pi\sigma_i^2} \int_{-\infty}^{+\infty} \left(\int_{\frac{1}{2}}^{\infty} e^{-(x-p_i)^2/2\sigma_i^2}\, dx\right) e^{-(p_i-\pi_i)^2/2\sigma_i^2}\, dp_i \\ &= \frac{1}{2\pi\sigma_i^2} \int_{-\infty}^{+\infty} \int_{\frac{1}{2}}^{\infty} e^{-[(x-p_i)^2+(p_i-\pi_i)^2]/2\sigma_i^2}\, dx\, dp_i . \end{aligned}$$

Now

$$(x - p_i)^2 + (p_i - \pi_i)^2 = \frac{(x - \pi_i)^2}{2} + 2\left(p_i - \frac{\pi_i + x}{2}\right)^2.$$

Hence

$$E(r_i) = \frac{1}{2\pi\sigma_i^2} \int_{\frac{1}{2}}^{\infty} e^{-(x-\pi_i)^2/4\sigma_i^2} \left(\int_{-\infty}^{+\infty} e^{-(p_i-\frac{1}{2}(\pi_i+x))^2/\sigma_i^2}\, dp_i\right) dx.$$

The second integral is equal to $\sqrt{\pi\sigma_i^2}$. Hence

$$E(r_i) = \frac{1}{2\sqrt{\pi\sigma_i^2}} \int_{\frac{1}{2}}^{\infty} e^{-(x-\pi_i)^2/4\sigma_i^2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{(\frac{1}{2}-\pi_i)/\sigma_i\sqrt{2}}^{\infty} e^{-x^2/2}\, dx .$$

---

[3] If π_i for any i were very close to $\frac{1}{2}$ then it would be of little use to poll the ith state. Hence, in this case formula (11) gives a small value for N_i. However, the π_i are never accurately known. The following procedure might be recommended for determining the best distribution of interviews: If for one particular i the sample value of π_i as estimated from a previous poll is too close to $\frac{1}{2}$ determine, using the N_i of the previous poll, that value $\bar{\pi}_i$ of π_i for which the probability is $\frac{3}{10}$ that p_i is larger than $\frac{1}{2}$ and substitute in (11) $\bar{\pi}_i$ for π_i. In all other cases substitute the sample value.

If several polls are taken it is advisable to use all of them but the last one to estimate as closely as possible the values of the π_i. The sample of the last poll before the election should be distributed according to (11).

From (12) we see that $E(r_i) < \rho_i$ if $\pi_i > \frac{1}{2}$ and $E(r_i) > \rho_i$ if $\pi_i < \frac{1}{2}$. Thus in every case this estimate is more biased than the estimate (1).

On the other hand, we shall now show that $E[(\epsilon_i - r_i)^2]$ is always smaller than $E[(\epsilon_i - e_i)^2]$. Since $\epsilon_i = 1$ if $\pi_i > \frac{1}{2}$ and $\epsilon_i = 0$ if $\pi_i < \frac{1}{2}$ it is easy to verify that $E[(\epsilon_i - r_i)^2]$ has the same value for $\pi_i = a$ as for $\pi_i = 1 - a$ and the same is true for $E[(\epsilon_i - e_i)^2]$. We may, therefore, without loss of generality assume that $\pi_i < \frac{1}{2}$.

Thus we have to show that

$$E(r_i^2) \leq E(e_i^2) = \rho_i = \int_{(\frac{1}{2}-\pi_i)/\sigma_i}^{\infty} e^{-\frac{1}{2}x^2}\,dx \qquad \text{if } \pi_i < \tfrac{1}{2}. \tag{13}$$

We have

$$E(r_i^2) = \frac{1}{\sqrt{2\pi}\sigma_i}\int_{-\infty}^{+\infty} e^{-(p_i-\pi_i)^2/2\sigma_i^2}\left(\int_{(\frac{1}{2}-p_i)/\sigma_i}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2}\,dx\right)^2 dp_i$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\int_{\frac{1}{2}}^{\infty}\int_{\frac{1}{2}}^{\infty} \frac{1}{2\pi\sigma_i^3}\, e^{-(1/2\sigma_i^2)Q(x,y,p_i)}\,dx\,dy\,dp_i\,.$$

Now

$$Q(x, y, p_i) = (x - p_i)^2 + (y - p_i)^2 + (p_i - \pi_i)^2$$

$$= 3\left(p_i - \frac{x + y + \pi_i}{3}\right)^2 + \frac{1}{6}(x + y - 2\pi_i)^2 + \frac{1}{2}(x - y)^2.$$

Putting

$$p_i' = \frac{\sqrt{3}\left(p_i - \dfrac{x + y + \pi_i}{3}\right)}{\sigma_i}, \quad x' = \frac{1}{\sqrt{6}}\,\frac{(x + y - 2\pi_i)}{\sigma_i},$$

$$y' = \frac{1}{\sqrt{2}}\,\frac{(x - y)}{\sigma_i}, \quad \frac{1 - 2\pi_i}{\sqrt{6}\sigma_i} = a,$$

we obtain

$$E(r_i^2) = \frac{1}{(\sqrt{2\pi})^3}\int_{-\infty}^{+\infty}\int_{a}^{\infty} e^{-\frac{1}{2}p'^2} e^{-\frac{1}{2}x'^2}\left(\int_{\sqrt{3}(a-x')}^{\sqrt{3}(x'-a)} e^{-\frac{1}{2}y'^2/2}\,dy'\right)dx'\,dp'$$

$$= \frac{1}{2\pi}\int_{a}^{\infty} e^{-\frac{1}{2}x^2}\int_{\sqrt{3}(a-x)}^{\sqrt{3}(x-a)} e^{-y^2}\,dy\,dx.$$

Now for $\pi_i = \frac{1}{2}$ we have $a = 0$, and for $\pi_i < \frac{1}{2}$ we have $a > 0$. For $a = 0$ we obviously have $E(r_i^2 \leq E(e_i^2)$. Further $\lim_{a\to\infty} E(r_i^2) = \lim_{a\to\infty} E(e_i^2) = 0$ hence (13) is proved if we can show that

$$F(a) = E(r_i^2) - E(e_i^2) = \frac{1}{2\pi}\int_{a}^{\infty} e^{-\frac{1}{2}x^2}\int_{\sqrt{3}(a-x)}^{\sqrt{3}(x-a)} e^{-\frac{1}{2}y^2}\,dy\,dx - \frac{1}{\sqrt{2\pi}}\int_{\sqrt{\frac{3}{2}}a}^{\infty} e^{-\frac{1}{2}x^2}\,dx$$

is a monotonically increasing function of a. Differentiating $F(a)$ with respect to a we obtain

$$
\begin{aligned}
\frac{dF(a)}{da} &= -\frac{\sqrt{3}}{\pi}\int_a^\infty e^{-\frac{1}{2}(4x^2-6ax+3a^2)} + \frac{\sqrt{3}}{2\sqrt{\pi}}\, e^{-(3/4)a^2} \\
&= \frac{-\sqrt{3}}{\pi}\, e^{-(3/4)a^2}\int_a^\infty e^{-4(x-(3/4)a)^2}\,dx + \frac{\sqrt{3}}{2\sqrt{\pi}}\, e^{-(3/4)a^2} \qquad (14)\\
&= \frac{-\sqrt{3}}{2\pi}\, e^{-(3/4)a^2}\int_{\frac{a}{2}}^\infty e^{-\frac{1}{2}x^2}\,dx + \frac{\sqrt{3}}{2\sqrt{\pi}}\, e^{-(3/4)a^2}.
\end{aligned}
$$

Hence for $a \geq 0$ we have

$$\frac{dF}{da} \geq \frac{-\sqrt{3}}{2\sqrt{2\pi}}\, e^{-\frac{3}{4}a^2} + \frac{\sqrt{3}}{2\sqrt{\pi}}\, e^{-\frac{3}{4}a^2} \geq 0.$$

Hence we have proved

$$
\begin{aligned}
E[(\epsilon_i - r_i)^2] &= \frac{1}{2\pi}\int_{|a|}^\infty e^{-\frac{1}{2}x^2}\int_{\sqrt{3}(|a|-x)}^{\sqrt{3}(x-|a|)} e^{-y^2}\,dy\,dx \leq E[(\epsilon_i - e_i)^2], \qquad (15)\\
a &= \frac{1-2\pi_i}{\sqrt{6}\,\sigma_i}.
\end{aligned}
$$

Since

$$E[(\epsilon_i - e_i)^2] - E[(\epsilon_i - r_i)^2]$$

is largest when $\pi_i = \frac{1}{2}$ we also have

$$E[(\epsilon_i - r_i)^2] \geq |\epsilon_i - \rho_i| - \left[\frac{1}{2} - \frac{1}{2\pi}\int_0^\infty e^{-\frac{1}{2}x^2}\int_{-\sqrt{3}x}^{+\sqrt{3}x} e^{-\frac{1}{2}y^2}\,dy\,dx\right]$$

or

$$|\epsilon_i - \rho_i| \geq E[(\epsilon_i - r_i)^2] \geq \frac{1}{2\pi}\int_0^\infty e^{-\frac{1}{2}x^2}\int_{-\sqrt{3}x}^{+\sqrt{3}x} e^{-\frac{1}{2}y^2}\,dy\,dx - |\tfrac{1}{2} - \rho_i|. \qquad (16)$$

Because of (15), r_i although more biased may in many cases be preferable to e_i as an estimate of ϵ_i.

NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

A COMBINATORIAL FORMULA AND ITS APPLICATION TO THE THEORY OF PROBABILITY OF ARBITRARY EVENTS[1]

By Kai-Lai Chung and Lietz C. Hsu

National Southwest Associated University, Kunming, China

An important principle, known as a proposition in formal logic or the method of cross-classification can be stated as follows.[1]

Let F and f be any two functions of combinations out of $(\nu) = (1, 2, \cdots, n)$. Then the two formulas

$$F((\alpha)) = \sum_{(\beta) \in (\nu)-(\alpha)} f((\alpha) + (\beta)) \tag{1.1}$$

$$f((\alpha)) = \sum_{(\beta) \in (\nu)-(\alpha)} (-1)^b F((\alpha) + (\beta)) \tag{2.1}$$

are equivalent.

As an immediate application to the theory of probability of arbitrary events, we have the set of inversion formulas[2]

$$p((\alpha)) = \sum_{(\beta) \in (\nu)-(\alpha)} p[(\alpha) + (\beta)] \tag{3.1}$$

$$p[(\alpha)] = \sum_{(\beta) \in (\nu)-(\alpha)} (-1)^b p((\alpha) + (\beta)) \tag{4.1}$$

where $p((\alpha))$ is the probability of the occurrence of at least $E_{\alpha_1}, E_{\alpha_2}, \cdots, E_{\alpha_a}$ out of n arbitrary events $E_1, E_2, \cdots, E_n$ and $p[(\alpha)]$ is the probability of the occurrence of $E\alpha_1, E\alpha_2, \cdots, E\alpha_a$ and no others among the n events, $(\alpha_1, \alpha_2, \cdots, \alpha_a)$ denoting a combination of the integers $(1, 2, \cdots, n)$. They can be made to play a central rôle in the theory, since they supply a method for converting the fundamental systems of probabilities, $p[(\alpha)]$ and $p((\alpha))$, one into the other.

We may further generalize (1.1) and (2.1) by considering combinations with repetitions. Let such a combination be written as

$$(\alpha) = (\alpha^r) = (\alpha_1^{r_1}\alpha_2^{r_2}\cdots\alpha_a^{r_a})$$

---

[1] For the notations and definitions see K. L. Chung, "On fundamental systems of probabilities of a finite number of events," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 123–133.

[2] Cf. Fréchet, *Les probabilités associées à un système d'événements compatibles et dépendants*, Hermann, Paris (1939), formulas (55) and (58).

where r_i ($r_i \geq 1$) denotes the number of repetitions of the number α_i, $i = 1, 2, \cdots, a$. Correspondingly we write

$$(\alpha)' = (\alpha_1\alpha_2 \cdots \alpha_a)$$

and call it the reduced combination corresponding to (α).

If there are n distinct elements $(1, 2, \cdots, n)$ in question, we may write every combination in the form

$$(1^{r_1}2^{r_2} \cdots n^{r_n})$$

where each r_i is zero or a positive integer. We say that $(1^{s_1}2^{s_2} \cdots n^{s_n})$ belongs to $(1^{r_1}2^{r_2} \cdots n^{r_n})$ and write

$$(1^{s_1}2^{s_2} \cdots n^{s_n}) \in (1^{r_1}2^{r_2} \cdots n^{r_n})$$

if and only if for each i, $i = 1, 2, \cdots, n$, we have $s_i \leq r_i$. We write

$$(1^{r_1}2^{r_2} \cdots n^{r_n}) + (1^{s_1}2^{s_2} \cdots n^{s_n}) = (1^{r_1+s_1}2^{r_2+s_2} \cdots n^{r_n+s_n});$$

and if $(1^{s_1}2^{s_2} \cdots n^{s_n}) \in (1^{r_1}2^{r_2} \cdots n^{r_n})$,

$$(1^{r_1}2^{r_2} \cdots n^{r_n}) - (1^{s_1}2^{s_2} \cdots n^{s_n}) = (1^{r_1-s_1}2^{r_2-s_2} \cdots n^{r_n-s_n}).$$

We define a generalized Möbius function $\mu((\alpha))$ for combinations (with or without repetitions) as follows

$$\mu((\alpha)) = \begin{matrix} (-1)^a & \text{if} & (\alpha) = (\alpha)' \\ 0 & \text{if} & (\alpha) \neq (\alpha)'. \end{matrix}$$

This function has the property

$$\sum_{(\beta) \in (\alpha)} \mu((\beta)) = \begin{matrix} 1 & \text{if} & (\alpha) = (0) \\ 0 & \text{if} & (\alpha) \neq (0). \end{matrix}$$

For we have

$$\sum_{(\beta) \in (\alpha)} \mu((\beta)) = \sum_{(\beta) \in (\alpha)'} (-1)^b = \sum_{b=0}^{a'} (-1)^b \binom{a'}{b}$$

$$= \begin{matrix} 1 & \text{if} & a' = 0 \\ 0 & \text{if} & a' \neq 0 \end{matrix} = \begin{matrix} 1 & \text{if} & (\alpha) = (0) \\ 0 & \text{if} & (\alpha) \neq (0). \end{matrix}$$

Now we state and prove the following general theorem.

THEOREM. *Let* $(\alpha)_i = (\alpha_{i1}^{r_{i1}}\alpha_{i2}^{r_{i2}} \cdots \alpha_{ia_i}^{r_{ia_i}})$ *and* $(\nu)_i = (1^{\lambda_{i1}}2^{\lambda_{i2}} \cdots n_i^{\lambda_{in_i}})$ *where* λ_{ij} *and* n_i *are finite and* $1 \leq r_{ij} \leq \lambda_{ij}$, $1 \leq a_i \leq n_i$ *for* $i = 1, 2, \cdots, m$ *and* $j = 1, 2, \cdots, n_i$. *Then for any two functions of the m combinations (with repetitions),* $(\alpha)_1, (\alpha)_2, \cdots, (\alpha)_m$ *out of* $(\nu)_1, (\nu)_2, \cdots, (\nu)_m$, *the two sets of formulas:*

$$F((\alpha)_1, (\alpha)_2, \cdots (\alpha)_m) = \sum_{(\beta)_i \in (\nu)_i - (\alpha)_i} f((\alpha)_1 + (\beta)_1, (\alpha)_2 + (\beta)_2, \cdots, (\alpha)_m + (\beta)_m) \tag{1}$$

and

$$f((\alpha)_1, (\alpha)_2, \cdots, (\alpha)_m) \tag{2}$$
$$= \sum_{(\beta)_i \in (\nu)_i - (\alpha)_i} \left[\prod_{i=1}^{m} \mu((\beta)_i)\right] F((\alpha)_1 + (\beta)_1, (\alpha)_2 + (\beta)_2, \cdots, (\alpha)_m + (\beta)_m)$$

are equivalent.

PROOF. To deduce (2) from (1)

$$\sum_{(\beta)_i \in (\nu)_i - (\alpha)_i} \left[\prod_{i=1}^{m} \mu((\beta)_i)\right] F((\alpha)_1 + (\beta)_1, \cdots, (\alpha)_m + (\beta)_m)$$
$$= \sum_{(\beta)_i \in (\nu)_i - (\alpha)_i} \left[\prod_{i=1}^{m} \mu((\beta)_i)\right] \sum_{(\gamma)_i \in (\nu)_i - (\alpha)_i - (\beta)_i} \cdot f((\alpha)_1 + (\beta)_1 + (\gamma)_1, \cdots, (\alpha)_m + (\beta)_m + (\gamma)_m)$$
$$= \sum_{(\delta)_i \in (\nu)_i - (\alpha)_i} f((\alpha)_1 + (\delta)_1, \cdots, (\alpha)_m + (\delta)_m) \cdot \sum_{(\gamma)_i \in (\delta)_i} \prod_{i=1}^{m} \mu((\delta)_i - (\gamma)_i).$$

Evidently we have

$$\sum_{(\gamma)_i \in (\delta)_i} \prod_{i=1}^{m} \mu((\delta)_i - (\gamma)_i) = \prod_{i=1}^{m} \left\{ \sum_{(\gamma)_i \in (\delta)_i} \mu((\delta)_i - (\gamma)_i) \right\}$$
$$= \prod_{i=1}^{m} \left\{ \sum_{(\gamma)_i \in (\delta)_i} \mu((\gamma)_i) \right\} = \begin{matrix} 1 & \text{if } (\delta)_i = (0) \text{ for } i = 1, \cdots, m \\ 0 & \text{otherwise} \end{matrix}$$

by the property of the μ-function. Hence the preceding sum reduces to $f((\alpha)_1, \cdots, (\alpha)_m)$ in accord with (2).

(1) is deduced from (2) in a similar way.

Although the general case is not without importance in the treatment of several sets of events,[3] we shall for the sake of convenience restrict ourselves to the special case $m = 1$.

In order to apply these formulas we must first introduce combinations with repetitions into the theory of arbitrary events. This can be done in various ways. Firstly, we may consider the number of occurrences of each event in a given time-interval or in a series of trials not necessarily independent. Secondly, we may regard each event as possessing various degrees of intensity. If the event E_i occurs r_i times in a given time-interval or occurs with r_i degrees of intensity, we write it as $E_i^{r_i}$. Hereafter we shall make use of the first interpreta-

[3] Cf. FRÉCHET, Loc. Cit. pp. 50–52; also, K. L. CHUNG, "Generalization of Poincaré's formula in the theory of probability," *Annals of Math. Stat.*, Vol. 14 (1943). We may note that our general theorem may be used to give another proof of the generalized Poincaré's formula for several sets of events.

tion and we shall assume that the maximum number of occurrences of each event is finite:

$$0 \leqq r_i \leqq \lambda_i, \qquad i = 1, \cdots, n.$$

We define

$p[E_1^{r_1} \cdots E_n^{r_n}] = p[(\nu^r)] =$ The probability that E_i occurs exactly r_i times in the given time-interval.

$p(E_1^{r_1} \cdots E_n^{r_n}) = p((\nu^r)) =$ The probability that E_i occurs at least r_i times in the given time-interval.

These quantities play the same rôle as the $p[(\alpha)]$'s and $p((\alpha))$'s in the ordinary theory. Evidently the probability of every complex event in question can be expressed as the sum of certain $p[(\nu^r)]$'s. To prove that the $p((\nu^r))$'s also form a fundamental system of quantities we have only to express $p[(\nu^r)]$'s in terms of the $p((\nu^r))$'s. This is given immediately by an application of the general theorem with $m = 1$. For we have in an obvious way

$$p(E_1^{r_1} \cdots E_n^{r_n}) = \sum_{r_i \leqq t_i \leqq \lambda_i} p[E_1^{t_1} \cdots E_n^{t_n}]$$

or

$$(3) \qquad p((\nu^r)) = \sum_{(\nu^s) \in (\nu^\lambda)-(\nu^r)} p[(\nu^r) + (\nu^s)] = \sum_{(\nu^s) \in (\nu^{\lambda-r})} p[(\nu^{r+s})].$$

Hence we obtain the inversion

$$(4) \qquad p[(\nu^r)] = \sum_{(\nu^s) \in (\nu^\lambda)-(\nu^r)} \mu((\nu^s))p((\nu^r) + (\nu^s)).$$

Let (α') denote a running combination without repetitions. Then since $\mu((\nu^s)) = 0$ unless (ν^s) is a (ν'),

$$(4') \quad p[(\nu^r)] = \sum_{(\alpha') \in (\nu^{\lambda-r})} \mu((\alpha'))p((\nu^r) + (\alpha')) = \sum_{(\alpha') \in (\nu^{\lambda-r})} (-1)^a p((\nu^r) + (\alpha'))^{\bullet}$$

The set of formulas (3) and (4) generalize (3.1) and (4.1).

Corresponding to the $p_{[a]}((\nu))$ for the ordinary events we define for $a + b + \cdots = n$ and $r, s, \cdots$ all distinct:

$p_{[a]^r,[b]^s,\cdots}(E_1^{\lambda_1} \cdots E_n^{\lambda_n}) =$ The probability that among n events $E_1, E_2, \cdots, E_n$ exactly a events occur r times, exactly b events occur s times and so on.

By (4) we easily obtain

$$(5) \qquad \begin{aligned} &p_{[a]^r,[b]^s,\cdots}((\nu^\lambda)) \\ &= \sum_S \sum_{(\nu^x) \in (\nu^\lambda)-((\alpha)^r+(\beta)^s+\cdots)} \mu((\nu^x))p((\nu^x) + (\alpha)^r + (\beta)^s + \cdots) \end{aligned}$$

where $(\alpha)^r = (E_{\alpha_1}^r \cdots E_{\alpha_a}^r)$, $(\beta)^s = (E_{\beta_1}^s \cdots E_{\beta_b}^s)$, $\cdots$ and the first summation is a symmetric sum which extends to all $n!/a!b! \cdots$ different combinations $(\alpha_1 \cdots \alpha_a)$, $(\beta_1 \cdots \beta_b)$, $\cdots$ out of $(\nu) = (1, 2 \cdots n)$.

The equality (5) is obviously a generalization of Poincaré's formula.

Similarly for the probabilities in the definition of which the word "exactly"

is sometimes substituted for the words "at least." Of course we can express all of them in terms of the $p[(\nu^r)]$'s or of the $p((\nu^r))$'s. However elegant formulas such as in the ordinary theory seem to be lacking.

Finally, we may also consider conditions of existence for the $p[(\nu^r)]$'s and the $p((\nu^r))$'s. For the former system the conditions are that they be all non-negative and that their sum be 1. For the latter system, the conditions are given by (4′), viz. for every $(\nu^r) \epsilon (\nu^\lambda)$,

$$\sum_{(\alpha') \epsilon (\nu^{\lambda - r})} \mu((\alpha'))p((\nu^r) + (\alpha)) \geqq 0.$$

These conditions are necessary and sufficient since (3) and (4) are equivalent.

ON THE MECHANICS OF CLASSIFICATION

By Carl F. Kossack

University of Oregon

1. Introduction. Wald[1] has recently determined the distribution of the statistic U to be used in the classification of an observation, z_i $(i = 1, 2, \cdots, p)$, as coming from one of two populations. He also determined the critical region which is most powerful for such a classification. It is the purpose of this paper to show how such a classification statistic under the assumption of large sampling can be applied in an actual problem and to present a systematic approach to the necessary computations.

The data used in this demonstration are those which were obtained from the A.S.T.P. pre-engineering trainees assigned to the University of Oregon. The problem considered is that of classifying a trainee as to whether he will do unsatisfactory or satisfactory work[2] in the first term mathematics course (Intermediate Algebra). The variables used in the classification are: (1) A Mathematics Placement Test Score. This is the score obtained by the trainee on a fifty-minute elementary mathematics test (including elementary algebra). The test was given to each trainee on the day that he arrived on the campus. (2) A High School Mathematics Score. A trainee's high school mathematics record was made into a score by giving 1 point to students who had had no high school algebra, 2 points to students with an F in first-year, high-school algbra and no second-year algebra, 3 points for a D, $\cdots$, 10 points for an average grade of A in first- and second-year algebra. (3) The Army General Classification Test Score. An individual needed a score of 115 or better in order to be assigned to the A.S.T.P. These data were obtained for 305 trainees along with the actual

[1] Abraham Wald, "On a statistical problem arising in the classification of an individual into one of two groups," *Annals of Math. Stat.*, Vol. 15, (1944), No. 2.

[2] Unsatisfactory work was defined as a grade of F or D in the course (failure or the lowest passing grade).

A.S.T.P. = Army Specialized Training Program

grade made by them in the algebra course. Trainees who had had college work were not included in the study.

2. Steps in the Computation of U and the Critical Region. Let

π_1 be the population of individuals who do unsatisfactory work in their first-term mathematics course.

π_2 be the population of individuals who do satisfactory work.

N_1 and N_2 = respectively the number of observed individuals in π_1 and π_2.

$x_{1\alpha}$ and $y_{1\alpha}$ = respectively the Mathematics Placement Test Score for the αth individual observed in π_1 and π_2.

$x_{2\alpha}$ and $y_{2\alpha}$ = respectively the High School Mathematics Score.

$x_{3\alpha}$ and $y_{3\alpha}$ = respectively the Army General Classification Test Score.

Step 1. *Computation of Summations*

$$
\begin{aligned}
N_1 &= 96 & N_2 &= 209 \\
\sum_\alpha x_{1\alpha} &= 3570 & \sum_\alpha y_{1\alpha} &= 11450 \\
\Sigma x_{2\alpha} &= 547 & \Sigma y_{2\alpha} &= 1567 \\
\Sigma x_{3\alpha} &= 11745 & \Sigma y_{3\alpha} &= 26684 \\
\Sigma x_{1\alpha}^2 &= 145476 & \Sigma y_{1\alpha}^2 &= 672452 \\
\Sigma x_{2\alpha}^2 &= 3509 & \Sigma y_{2\alpha}^2 &= 12577 \\
\Sigma x_{3\alpha}^2 &= 1439559 & \Sigma y_{3\alpha}^2 &= 3421996 \\
\Sigma x_{1\alpha}x_{2\alpha} &= 21012 & \Sigma y_{1\alpha}y_{2\alpha} &= 88774 \\
\Sigma x_{1\alpha}x_{3\alpha} &= 436964 & \Sigma y_{1\alpha}y_{3\alpha} &= 1469302 \\
\Sigma x_{2\alpha}x_{3\alpha} &= 66731 & \Sigma y_{2\alpha}y_{3\alpha} &= 200150 \\
\Sigma(x_{1\alpha} - \bar{x}_1)^2 &= 12716.625 & \Sigma(y_{1\alpha} - \bar{y}_1)^2 &= 45167.311 \\
\Sigma(x_{2\alpha} - \bar{x}_2)^2 &= 392.240 & \Sigma(y_{2\alpha} - \bar{y}_2)^2 &= 828.249 \\
\Sigma(x_{3\alpha} - \bar{x}_3)^2 &= 2631.656 & \Sigma(y_{3\alpha} - \bar{y}_3)^2 &= 15125.876 \\
\Sigma(x_{1\alpha} - \bar{x}_1)(x_{2\alpha} - \bar{x}_2) &= 670.438 & \Sigma(y_{1\alpha} - \bar{y}_1)(y_{2\alpha} - \bar{y}_2) &= 2926.392 \\
\Sigma(x_{1\alpha} - \bar{x}_1)(x_{3\alpha} - \bar{x}_3) &= 196.812 & \Sigma(y_{1\alpha} - \bar{y}_1)(y_{3\alpha} - \bar{y}_3) &= 7427.359 \\
\Sigma(x_{2\alpha} - \bar{x}_2)(x_{3\alpha} - \bar{x}_3) &= -191.031 & \Sigma(y_{2\alpha} - \bar{y}_2)(y_{3\alpha} - \bar{y}_3) &= 83.837
\end{aligned}
$$

Step 2. *Computation of Statistics.*

$$
\begin{aligned}
\bar{x}_1 &= 37.188 & \bar{y}_1 &= 54.785 \\
\bar{x}_2 &= 5.6979 & \bar{y}_2 &= 7.4976 \\
\bar{x}_3 &= 122.3438 & \bar{y}_3 &= 127.6746
\end{aligned}
$$

$$
s_{ij} = \frac{\Sigma(x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) + \Sigma(y_{i\alpha} - y_i)(y_{j\alpha} - y_j)}{N_1 + N_2 - 2}
$$

$$
\begin{aligned}
s_{11} &= 191.04 & s_{12} &= 11.871 \\
s_{22} &= 4.0280 & s_{13} &= 25.162 \\
s_{33} &= 58.606 & s_{23} &= -.35378
\end{aligned}
$$

Step 3. *Computation of Inverse Matrix* $| s^{ij} |$

$$| s_{ij} | = \begin{vmatrix} 191.04 & 11.871 & 25.162 \\ 11.871 & 4.0280 & -.35378 \\ 25.162 & -.35378 & 58.606 \end{vmatrix} = 34053$$

$$\begin{array}{ll} s^{11} = .0069286 & s^{12} = -.020692 \\ s^{22} = .31019 & s^{13} = -.0030996 \\ s^{33} = .018459 & s^{23} = .010756 \end{array}$$

Step 4. *Computation of the Classification Equation.*

$$\begin{aligned} U = {} & [s^{11}(\bar{y}_1 - \bar{x}_1) + s^{12}(\bar{y}_2 - \bar{x}_2) + s^{13}(\bar{y}_3 - \bar{x}_3)]\cdot z_1 \\ & + [s^{21}(\bar{y}_1 - \bar{x}_1) + s^{22}(\bar{y}_2 - \bar{x}_2) + s^{23}(\bar{y}_3 - \bar{x}_3)]\cdot z_2 \\ & + [s^{31}(\bar{y}_1 - \bar{x}_1) + s^{32}(\bar{y}_2 - \bar{x}_2) + s^{33}(\bar{y}_3 - \bar{x}_3)]\cdot z_3 \end{aligned}$$

where z_i plays the same role for individuals to be classified as $x_{i\alpha}$ and $y_{i\alpha}$ do for observed individuals.

$$U = .068160\, z_1 + .25147\, z_2 + .063215\, z_3$$

Step 5. *Computation of the Critical Region (assuming* $W_1 = W_2$*)*

$$\bar{\alpha}_1 = .068160\, \bar{x}_1 + .25147\, \bar{x}_2 + .063215\, \bar{x}_3 = 11.702$$

$$\bar{\alpha}_2 = .068160\, \bar{y}_1 + .25147\, \bar{y}_2 + .063215\, \bar{y}_3 = 13.691$$

$$\tfrac{1}{2}(\bar{\alpha}_1 + \bar{\alpha}_2) = 12.696$$

Therefore,

For $U \leq 12.696$ classify the individual as coming from π_1 population.
For $U > 12.696$ classify the individual as coming from π_2 population.

Step 6. *Computation of the Efficiency of Classification.*

$$\begin{aligned} \bar{\sigma}^2 = {} & s^{11}(\bar{y}_1 - \bar{x}_1)(\bar{y}_1 - \bar{x}_1) + s^{12}(\bar{y}_1 - \bar{x}_1)(\bar{y}_2 - \bar{x}_2) + s^{13}(\bar{y}_1 - \bar{x}_1)(\bar{y}_3 - \bar{x}_3) \\ & + s^{21}(\bar{y}_2 - \bar{x}_2)(\bar{y}_1 - \bar{x}_1) + s^{22}(\bar{y}_2 - \bar{x}_2)(\bar{y}_2 - \bar{x}_2) + s^{23}(\bar{y}_2 - \bar{x}_2)(\bar{y}_3 - \bar{x}_3) \\ & + s^{31}(\bar{y}_3 - \bar{x}_3)(\bar{y}_1 - \bar{x}_1) + s^{32}(\bar{y}_3 - \bar{x}_3)(\bar{y}_2 - \bar{x}_2) + s^{33}(\bar{y}_3 - \bar{x}_3)(\bar{y}_3 - \bar{x}_3) \\ = {} & 1.5764. \end{aligned}$$

$$\frac{\bar{\alpha}_2 - \bar{\alpha}_1}{2\bar{\sigma}} = .792$$

$$P_1 = 1 - P_2 = \frac{1}{\sqrt{2\pi}} \int_{.792}^{\infty} e^{-t^2/2} = .2062$$

where P_1 is the probability of making an error of Type I, that is, of classifying an individual as one who will do satisfactory work when he actually does unsatisfactory work; and $1 - P_2$ is the probability of making an error of Type II,

that is, of classifying a student as one who will do unsatisfactory work when he actually does satisfactory work.

3. Conclusions. In using the above classification equation to classify the 305 trainees used in this study, 21 errors of Type I were made or 22.9 percent, while 50 errors of Type II were made or 23.9 percent. These percentages seem reasonably close to the expected 20.6 percent.

NOTE ON AN IDENTITY IN THE INCOMPLETE BETA FUNCTION

By T. A. Bancroft

Iowa State College

Since the incomplete beta function has proved of some importance in statistics, it would appear that any additional information concerning its properties might at some time prove useful. In a paper by the author, [1], two identities in the incomplete beta function were incidentally obtained. They are as follows:

$$(p + q)I_x(p, q) = pI_x(p + 1, q) + qI_x(p, q + 1) \tag{1}$$

and

$$\begin{aligned}(p + q + 1)^{[2]}I_x(p, q) = (p + 1)^{[2]}I_x(p + 2, q) &+ 2pqI_x(p + 1, q + 1) \\ &+ (p + 1)^{[2]}I_x(p, q + 2),\end{aligned} \tag{2}$$

where the incomplete beta function $I_x(p, q) = \dfrac{B_x(p, q)}{B(p, q)}$, etc., and $(p + 1)^{[2]}$, etc. refer to the standard factorial notation.

Written in the above form these two identities suggest a possible general identity to which they belong as special cases. The third special case suggested is:

$$\begin{aligned}(p + q + 2)^{[3]}I_x(p, q) &= (p + 2)^{[3]}I_x(p + 3, q) \\ &+ 3(p + 1)^{[2]}qI_x(p + 2, q + 1) + 3p(q + 1)^{[2]}I_x(p + 1, q + 2) \\ &+ (q + 2)^{[3]}I_x(p, q + 3).\end{aligned} \tag{3}$$

The general formula suggested is

$$\begin{aligned}(p + q + n - 1)^{[n]} I_x (p, q) = \sum_{r=0}^{n} \binom{n}{r} &(p + n - r - 1)^{[n-r]} \\ &\cdot (q + r - 1)^{[r]} I_x (p + n - r, q + r).\end{aligned} \tag{4}$$

To prove the general formula we write (4) as

$$\begin{aligned}(p + q + n - 1)^{[n]} I_x (p, q) = \sum_{r=0}^{n} \binom{n}{r} &(p + n - r - 1)^{[n-r]} \\ &\cdot (q + r - 1)^{[r]} \frac{B_x (p + n - r, q + r)}{B(p + n - r, q + r)}.\end{aligned} \tag{5}$$

By expanding and simplifying it is easy to show that

$$\frac{(p + n - r - 1)^{[n-r]} (q + r - 1)^{[r]}}{\mathrm{B}(p + n - r, q + r)} = \frac{(p + q + n - 1)^{[n]}}{\mathrm{B}(p, q)}. \tag{6}$$

Using (6) the right hand side of (5) reduces to

$$\frac{(p + q + n - 1)^{[n]}}{\mathrm{B}(p, q)} \sum_{r=0}^{n} \binom{n}{r} \mathrm{B}_x (p + n - r, q + r). \tag{7}$$

The summed function in (7) reduces to

$$\int_0^x x^{p-1} (1 - x)^{q-1} [x + (1 - x)]^n \, dx = \mathrm{B}_x (p, q), \tag{8}$$

which proves the identity.

Although the general identity is quite simple to prove, it does not seem to have appeared in the literature.

REFERENCE

[1] BANCROFT, T. A. "On biases in estimation due to the use of preliminary tests of significance," *Annals of Math. Stat.*, Vol. 15 (1944), No. 2.

NEWS AND NOTICES

Readers are invited to submit to the Secretary of the Institute news items of interest

Personal Items

Archie Blake is now employed as a ballistician with the Ballistic Research Laboratory at Aberdeen Proving Ground.

Robert V. Bonnar is now employed as Associate Technologist at the Mare Island Navy Yard.

Professor W. G. Cochran has returned to his regular duties at Iowa State College.

Mrs. Bianca Cody (Bianca Rivoli) is now Statistician for the James O. Peck Research Company, 12 East 41st Street, New York City.

Associate Professor William Feller of Brown University has been appointed Professor of Mathematics at Cornell University.

Professor John Kenney of the University of Wisconsin is now located at the Milwaukee branch of the University.

Myra Levine is now Assistant Mathematical Statistician with the Statistical Research Group at Columbia University.

Mrs. Harold Michaelis (Ruth E. Jolliffe) is 5th Naval District Statistician at the Naval Operating base in Norfolk, Va.

Emma Spaney is Statistician for the Committee on Measurement of the National League of Nursing Education.

Professor J. A. Shohat of the University of Pennsylvania died October 8, 1944.

Mr. Redford T. Webster of the Western Electric Company died July 31, 1944.

New Members

The following persons have been elected to membership in the Institute:

Boddie, John B., Jr. Chief, Program Section, Budget Division, Washington, D. C. *2628 Tunlaw Road, N.W.*

Bruner, Nancy M.A. (Iowa) Statistician, Western Auto Supply Co., Kansas City, Mo. *7511 Main St.*

Christopher, Edward E. B.S. (Mass. Inst. Tech.) Statistician, Signal Corps. *5704 North 26th St., Arlington, Va.*

Cowden, Dudley J. Ph.D. (Columbia) Prof. of Economics, Univ. of North Carolina. *Box 515, Chapel Hill, North Carolina.*

Cynamon, Manuel M.S. (City Coll., N. Y.) Personnel Tech., Personnel Res. Sec., Adj. General's Office, War Dept. *10 Ave. P, Brooklyn 4, N. Y.*

Evensen, Edward J. On military leave from Metropolitan Life Ins. Co. (Actuarial Sec.) *Sv. Co., 1st Sp. Sv. Force.*

Green, Earl L. Ph.D. (Brown) 1st Lieut., A.C., Chief, Dept. of Statistics. AAF School of Aviation Medicine, Randolph Field, Texas.

Groves, William Brewster B.S. (Antioch) Economist, Off. of Price Administration. *520 Decatur St., N.W., Washington, D. C.*

Hornseth, Richard Allen M.A. (Wisconsin) Res. Assistant in Sociology, Univ. of Wisconsin. *207 N. Randall, Madison 5, Wis.*

Kinsler, David M. M.A. (Chicago) Chief, Analytical Section, Arms & Ammunition Division, Aberdeen Proving Ground, Maryland.

Kopp, Paul J. M.A. (Duke) Major, Chemical Warfare Service, U. S. A. *1305 North Adams St., Arlington, Va.*

Massey, Frank Jones, Jr. M.A. (California) Associate, Dept. of Math., Univ. of California, Berkeley, Calif. *1364 Union St., San Francisco 9, Calif.*

Orcutt, Guy H. Ph.D. (Michigan) Instr. Economics Dept., Mass. Inst. of Tech., Cambridge, Mass.

Rakesky, Sophie M.S. (Michigan) Statistician, W. K. Kellogg Foundation, Battle Creek, Mich.

Roberts, Jean M.S. (Minnesota) Statistician, Child Welfare Res. Analyst. *929 Goodrich Ave., St. Paul 5, Minn.*

Schietroma, William B.S.S. (Coll. of City of N. Y.) Research Assistant. *316 East 116th St., New York, N. Y.*

Schlorek, Mary A. A.B. (Adelphi) Research Statistician, National Broadcasting Co., 30 Rockefeller Plaza, New York, N. Y.

deSousa, Alvaro Pedro B.E. (Liverpool) Vice-Governor, Banco de Portugal. *Monserrate, Rua Infante de Sagres, Estoril, Portugal.*

Steele, Floyd George M.S. (Calif. Inst. of Tech.) Stat. Analyst, Douglas Aircraft. *18168 Roosevelt Highway, Pacific Palisades, Calif.*

Thom, Herbert C. S. 6130 18th Rd., N., Arlington, Va.

Report of the Fifth Pittsburgh Chapter Meeting

The fifth meeting of the Pittsburgh Chapter of the Institute of Mathematical Statistics was held at Engineering Hall, Carnegie Institute of Technology on Saturday, November 25, 1944. The meeting was held as a joint session with the Pittsburgh Quality Control Society. Thirty-one persons attended the meeting, including the following six members of the Institute:

George Eldredge, H. J. Hand, C. R. Mummery, E. G. Olds, E. M. Schrock, J. V. Sturtevant.

The following papers were presented, with Mr. J. V. Sturtevant, of the Carnegie Illinois Steel Corporation, acting as chairman:

1. *Modified Application of Control Chart to the Use of Gauges on Machine Tool Work.*
   Dr. E. G. Olds, War Production Board, Washington, D. C.
2. *Application of Control Charts to Infrequent Inspection of Machine Operations.*
   W. D. Angst, Thompson Aircraft Products Company, Cleveland, Ohio.
3. *Application of Control Chart Techniques to Checking Reproducibility of Chemical Analysis.*
   H. A. Stobbs, Wheeling Steel Corporation, Steubenville, Ohio.
4. *Statistical Principles of Experimental Design as Applied to Tests Conducted in Manufacturing Operations.*
   Dr. B. Epstein, Westinghouse Electric & Manufacturing Co., East Pittsburgh, Pa.

H. J. Hand,
Secretary-Treasurer, Pittsburgh Chapter

Educational Meetings of the Pittsburgh Chapter

The first of a series of educational meetings on methods of statistical computations given by the Pittsburgh Chapter was held on Saturday afternoon, January 20, 1945. Thirty-three persons attended the meeting, including the following three members of the Institute:

Thomas A. Elkins, H. J. Hand, J. V. Sturtevant.

The following program was presented:

1. *Potential Field for Industrial Applications of Statistical Method.*
   H. J. Hand, National Tube Company, Pittsburgh, Pa.
2. *Computations for Analysis of Variance and Experimental Design.*
   Ben Epstein, Westinghouse Electric & Manufacturing Company, East Pittsburgh, Pa.

It is planned to hold these meetings bi-weekly, on Saturday afternoons for an indefinite period in the future. Topics to be considered in the series will include:

1. Analysis of variance and covariance.
2. Design of experiments.
3. Tests of significance.
4. Probability and probability distributions.
5. Correlation and regression analysis, including the orthogonal coordinate method.
6. Tests of increased severity.
7. Sampling theory, including stratification.
8. Acceptance-rejection mathematics, Dodge sampling inspection tables.
9. Shewhart control chart techniques.
10. Analysis of runs.
11. Cycle analysis.
12. Factor analysis.

H. J. Hand,
Secretary-Treasurer, Pittsburgh Chapter

ANNUAL REPORT OF THE PRESIDENT OF THE INSTITUTE

Continuing the established tradition, the annual summer meeting was held at Wellesley, Massachusetts, August 12–13, 1944 in conjunction with the Summer Meetings of the American Mathematical Society and the Mathematical Association of America. A regional meeting was held in Washington, May 6–7, in conjunction with the meeting of the Washington Chapter of the American Statistical Association. The programs were arranged by the Program Committee: W. Feller, Chairman, W. G. Madow, and A. Wald.

Even though, under present war conditions, research in the field of probability and statistics is very much curtailed, enough papers in mathematical statistics of satisfactory quality have been proposed for publication in the *Annals* in 1944 to keep the total volume of material at approximately five hundred pages or the level of the last few years. However, the outlook for a sufficient number of satisfactory papers to maintain the usual volume of publication during 1945 does not look quite so favorable.

Looking into the future, the Institute must continue to furnish, through the *Annals*, a medium for the publication of all important results of original research in the field of mathematical statistics as they become available. To do otherwise would be suicide. At the same time we must take account of the growing need for comprehensive surveys of statistical theory on the part of other scientists, including not only social scientists but also physicists, chemists, biologists, and research engineers, whose interest in the contributions of mathematical statistics has been greatly stimulated during the war. Only the mathematical statiscian of broad competence can provide adequate critical surveys of this character. Perhaps some of this need can be met through survey articles published in the *Annals*, although it is not an easy matter to get capable men to do such work. Perhaps the time is not far off when the Institute must stimulate the preparation of such material by instituting an annual series of Colloquium Lectures patterned somewhat after those of the Mathematical Society, which could be published separately.

This is but one of many problems that the Institute faces in its post-war development. Not only must it assume the responsibility of stimulating and encouraging research and of publishing the results; it must also consider the problem of training the research statistician of tomorrow as well as those who are to apply mathematical statistics in the many fields of science. It also must assume some responsibility for keeping in contact with other scientists in order that the mathematical statistician may become acquainted with the unsolved statistical problems of the scientist. There are also many problems of a professional character that face the mathematical statistician in the future if he is to succeed in developing the profession of mathematical statistics to the level attained by some of the older scientific professions.

With the realization of the need for a concerted attack on some of these

problems, the Board of Directors at its meeting in May set up two committees, one on Training and Placement of Statisticians under Harold Hotelling and the other on Post-War Development of the Institute under W. G. Cochran. Interim reports received by the Board from both committees indicate that considerable progress has been made to date. They also indicate, however, that much more work remains to be done.

At the same meeting of the Board, a Budget and Finance Committee was set up, consisting of P. S. Dwyer, Chairman, C. H. Fischer, A. C. Olshen, and C. F. Roos, to prepare a report on the policy that should be followed by the Institute in respect to such items as investment of funds, advertising, preparation of an annual statement, and the like. Some of the work of this committee has already borne fruit, as, for example, in providing the actuarial basis for life membership adopted at the Wellesley meeting and in establishing certain principles to be used in conducting the business of the Institute.

A report of the Committee on Membership, W. G. Cochran, Chairman, P. S. Dwyer, and T. Koopmans, appears elsewhere in this issue of the *Annals*. Upon recommendation of this committee, the Board of Directors elected nine new fellows: Walter Bartky, C. I. Bliss, Gertrude M. Cox, P. A. Horst, M. G. Kendall, H. B. Mann, E. S. Pearson, Henry Scheffé, and W. A. Wallis.

The nominating committee for the year consisted of John Curtiss, Chairman, E. G. Olds, and F. F. Stephan. G. W. Snedecor served the Institute again as its representative on the Council of the A.A.A.S.

The annual election of the Institute just concluded by mail ballot resulted in the election of the following officers for 1945: W. E. Deming, President; W. G. Cochran, and J. L. Doob, Vice-Presidents.

WALTER A. SHEWHART
President, 1944

February 10, 1945

ANNUAL REPORT OF THE SECRETARY-TREASURER OF THE INSTITUTE

Accounts of the 1944 meetings of the Institute—the Wellesley meeting, the Washington regional meeting, and the Pittsburgh chapter meetings—have appeared in appropriate issues of the *Annals*.

At the Wellesley meeting a number of amendments to the Constitution and By-Laws were passed. These were published in the September, 1944, issue of the *Annals*. (The amended Constitution and By-Laws appear elsewhere in this issue.)

Due to a large extent to the cooperation of the membership in sending in nominations, the Institute enjoyed a large increase in membership during the year. There were some resignations and it was necessary to suspend fifteen persons at the end of 1944 because of failure to pay dues. It is apparent that, in some of these cases at least, our mail is not being received. Undoubtedly some of these memberships will be restored when contact is again established. As of January 1, 1945, there were 606 members, a net gain of approximately one hundred members.

During the year the Institute received gifts from Professor Harry Carver in the form of exchanges for early issues of the *Annals*, reprints of early articles, etc.

The Secretary-Treasurer wishes to acknowledge the continued assistance of Professor Lloyd Knowler in looking after the back issues of the *Annals* which are stored at Iowa City.

The following financial statement covers the period from December 22, 1943 to December 31, 1944 (the books and records of the Treasurer have been audited by Professor Thomas A. Bickerstaff and were found to be in agreement with the statement as submitted):

FINANCIAL STATEMENT

December 22, 1943, to December 31, 1944

RECEIPTS

BALANCE ON HAND, DECEMBER 22, 1943		$3,715.05
DUES		
1944 and before	$2,995.31	
1945 and 1946	1,127.00	
Life	330.00	
		4,452.31
SUBSCRIPTIONS		
1944 and before	$1,301.94	
1945 and 1946	883.94	
		2,185.88
SALE OF BACK NUMBERS		1,385.02
MISCELLANEOUS		6.15
Total Receipts		$11,744.41

EXPENDITURE

ANNALS—CURRENT		
Office of Editor	$273.77	
Waverly Press	3,448.51	
		3,722.28
ANNALS—BACK NUMBERS		
Purchase from H. C. Carver	$149.40	
Iowa City Office	96.26	
		245.66
OFFICE OF SECRETARY-TREASURER		
Printing, mimeographing, programs, etc. (including stamped envelopes)	$377.00	
Postage and supplies	68.02	
Clerical help	455.94	
Moving office from Pittsburgh	55.79	
		956.75
MISCELLANEOUS		29.07
BALANCE ON HAND, DECEMBER 31, 1944		6,790.65
		$11,744.41

No unpaid bills were in the hands of the Treasurer as of December 31, 1944, and aside from an additional $100.00 which the Board has designated for *Annals* expense for 1944, there were no large bills outstanding.

Accounts receivable as of December 31, 1944, amounted to $303.73. Many of these accounts are current accounts while some of the older ones are accounts with firms in India, which probably will be collected eventually.

The American Library Association continued with its purchase of thirty sets of Volume XV of the *Annals* (for post war distribution) and the Universal Trading Corporation (representing the Chinese Government) purchased twenty sets of Volumes 11–17 inclusive. These orders contributed in no small way to the total 1944 income of $8,029.36.

The 1944 balance $6,790.65 (consisting of bank balance of $3,790.65 and $3,000.00 in government bonds) is $3,075.60 higher than it was on December 21, 1943. This increase is due in part to 1944 business and in part to the fact that unusually large payments toward future business, such as the $330.00 in life payments and the $1,127.00 in 1945 and 1946 dues, have been made.

To summarize the situation briefly, the Institute's 1944 activity has resulted in a gain of approximately $1,500.00 and we are about this much in advance of our usual position with reference to the payments of following years.

PAUL S. DWYER
Secretary-Treasurer.

December 31, 1944

REPORT OF THE MEMBERSHIP COMMITTEE OF THE INSTITUTE

Since the duties of this Committee are not defined in detail in the Constitution, the Board of Directors asked the Committee to prepare a statement describing the appropriate composition and function of the Committee on Membership. This work resulted in the preparation of amendments to the Constitution and By-laws. These amendments were passed at the business meeting at Wellesley College on August 13, 1944, and are printed in full in the September, 1944, issue of the *Annals* (p. 340).

In brief, the duties of the Committee are specified as follows in these amendments:

(a) The Committee holds the power of election to the grades of Member and Junior Member and makes recommendations to the Board of Directors with reference to placing members in the other grades of membership.

(b) It is the duty of the Committee to prepare and make available through the Secretary-Treasurer an announcement of the qualifications necessary for the different grades of membership and to review these qualifications periodically.

(c) The Committee considers plans for increasing the number of applicants for membership.

As permitted by the amendments referred to above, the power of election to the grades of Member and Junior Member was delegated by the Committee in August, 1944, to the Secretary-Treasurer, subject to certain reservations. The statement of qualifications for the different grades of membership as mentioned in (b) above is published below. At the August 13 meeting of the Board of Directors it was decided that no elections should be made at present to the grades of Honorary Member and Sustaining Member.

On the recommendation of the Membership Committee the following members were elected as Fellows by the Board of Directors: W. Bartky, C. I. Bliss, G. M. Cox, P. A. Horst, M. G. Kendall, H. B. Mann, E. S. Pearson, H. Scheffé, W. A. Wallis.

Statement of Qualifications for the Different Grades of Membership in the Institute of Mathematical Statistics

Member. The candidate shall either (a) be actively engaged in or show a serious interest in mathematical statistics, or (b) be interested in some applied field of statistics, with a desire to keep himself informed regarding recent developments in mathematical theory and techniques.

Junior Member.

1. Any undergraduate student of a collegiate institution is eligible for election as a Junior Member of the Institute of Mathematical Statistics provided that he or she is sponsored by a member of the Institute.
2. The annual dues ($2.50) must be submitted with the application.

3. Annual membership shall coincide with the calendar year and the Junior Member shall receive a complete volume of the *Annals of Mathematical Statistics* for the year in which he or she is elected.

4. Junior Membership shall be limited to a term of two years, but a Junior Member may apply for transfer to ordinary membership at the beginning of his second year.

Fellow.

1. The candidate shall have evidenced continuing activity in research in mathematical statistics by publication beyond his doctor's dissertation of independent work of merit. Normally two or three worthwhile papers beyond the dissertation will be required to establish this fact.

2. The first qualification may be partly or wholly waived in the case of (a) a candidate of well-established leadership among mathematical statisticians whose contributions to the development of the field of mathematical statistics other than sufficient published original research shall be judged of equal value or (b) a candidate of well-established leadership in the applications of mathematical statistics, whose work has contributed greatly to the utility of and the appreciation for mathematical statistics.

Honorary Member. A person of exceptional ability and acknowledged leadership in the field of mathematical statistics may be elected to the grade of Honorary Member by the Board of Directors, upon the recommendation of the Committee on Membership.

Sustaining Member. The Board of Directors shall have the power to elect to Sustaining Membership any individual, group or corporation that is interested in furthering the purposes for which the Institute was formed.

W. G. Cochran (*Chairman*)
W. E. Deming
P. S. Dwyer
T. Koopmans

February 10, 1945

PROGRESS REPORT OF THE COMMITTEE ON POST-WAR DEVELOPMENT OF THE INSTITUTE

In considering the post-war development of the Institute of Mathematical Statistics, the Committee has recognized two general problems:

A. The problem of what additional activities the Institute should undertake in order to provide further stimulus to the development of the field of mathematical statistics.

B. The problem of determining how the Institute can cooperate more effectively with the users of statistical techniques.

Because of rapidly increasing interest in the application of statistical methods in many different fields, the Committee has directed most of its attention thus far to Problem B; the present progress report is concerned with the work of the Committee on this problem. The Committee hopes to submit a report on Problem A at the end of 1945.

With respect to Problem B, it is the opinion of the Committee that a central organization for the statistical societies should be of common interest. Accordingly, a plan was worked out and submitted to the Board of Directors of the Institute at the Wellesley meeting of the Institute. This proposal and its present status are discussed below.

We believe that there is much to be gained from an organization that would form a link between the various statistical societies, and would have the following principal aims:

(1) To represent the members of the societies in all matters of common interest.

(2) To promote cooperation between statisticians working in the different fields of application, and between mathematical statistics, applied statistics, scientific research and the industries.

(3) To develop amongst the public an appreciation of the value of the statistical method in scientific inquiry.

It is our opinion that an organization similar to that of the Institute of Physics would be suitable. The statistical societies, while retaining their present autonomies, would become founding members of a corporation whose governing board would contain representatives from each society. In pursuance of its aims as outlined above, the new organization might:

(a) Take the lead in formulating policies on questions which concern all statisticians.

(b) Publish a journal of general interest to statisticians and undertake the routine work in connection with the publication of the journals of the individual societies, the societies retaining in full their present responsibility for the contents of their journals.

(c) Arrange joint meetings between different statistical societies and between statistical and other scientific societies.

(d) Assist new groups in organizing for their benefit, either under the auspices of one of the present societies or in a new society, which might at first be given associate membership and later full membership of the central organization.
(e) Take steps to bring news about the use of statistics in scientific research to the attention of the public and more particularly of leaders in industry, in federal, state and local agencies and in education.
(f) Investigate the demands for various types and degrees of statistical training, outline courses of training in statistics suitable for meeting these demands and make strenuous efforts to have the recommended courses of training put into effect, in order that statisticians can be of fullest service in the nation's work. In this connection an information and placement bureau may be an appropriate auxiliary.
(g) Institute an abstracting service in statistical methodology. This might take the form of a periodical publication of abstracts of papers with respect to their methodological content rather than their subject matter. The coverage would include journals of business, marketing, engineering, medicine and agriculture as well as purely statistical publications.

The financial needs of the new organization, which would maintain a paid full-time staff, may be met initially by contributions from the present societies. In view of the extra services which would be rendered to statisticians, some increase in the subscription rates of the present societies appears reasonable. A member who belongs to more than one of the present societies would pay the extra amount only once. Supplementary income might be derived from advertising in the journal of the central organization and from the establishment of sustaining or corporate memberships in the central organization.

At the time of the Wellesley meeting of the Board, there had been only informal contacts between members of this Committee and members of other statistical societies. We considered it our first task to obtain some consensus of opinion from the standpoint of the Institute of Mathematical Statistics. Following general approval by the Board of Directors of the Institute, members of the Committee discussed the proposal for a central organization with representatives of several other statistical societies. The American Statistical Association has a Committee to consider the future structure of the Association and this Committee brought the Institute proposal before the Board of Directors of the Association for action. As the oldest of the statistical societies, the American Statistical Association then invited participation in an intersociety committee by the Institute and nine other societies or sections, directly or indirectly concerned with statistical method. This committee is to explore the possibilities of coordinating the activities of the several statistical societies and report its recommendations back to each organization. The representatives have now been named and the first meeting was held on February 10, 1945, in New York. At this meeting the Institute was represented by W. G. Cochran and Lt. John H. Curtiss.

With regard to the problem of what additional activities the Institute should undertake in order to furnish additional stimulation to the development of the field of mathematical statistics, the Committee has discussed several ideas which appear promising. It is hoped to present a complete report on this phase of the Committee's work at the end of this year.

C. I. Bliss
W. G. Cochran (*Chairman*)
W. E. Deming
P. S. Olmstead
S. S. Wilks

February 12, 1945

CONSTITUTION
OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

ARTICLE I

Name and Purpose

1. This organization shall be known as the Institute of Mathematical Statistics.
2. Its object shall be to promote the interests of mathematical statistics.

ARTICLE II

Membership

1. The membership of the Institute shall consist of Members, Junior Members, Fellows, Honorary Members, and Sustaining Members.

2. Voting members of the Institute shall be (a) the Fellows, and (b) all others, Junior Members excepted, who have been members for twenty-three months prior to the date of voting.

3. No person shall be a Junior Member of the Institute for more than a limited term as determined by the Committee on Membership and approved by the Board of Directors.

ARTICLE III

Officers, Board of Directors, and Committee on Membership

1. The Officers of the Institute shall be a President, two Vice-Presidents, and a Secretary-Treasurer. The terms of office of the President and Vice-Presidents shall be one year and that of the Secretary-Treasurer three years. Elections shall be by majority ballots at Annual Meetings of the Institute. Voting may be in person or by mail.

(a) Exception. The first group of Officers shall be elected by a majority vote of the individuals present at the organization meeting, and shall serve until December 31, 1936.

2. The Board of Directors of the Institute shall consist of the Officers, the two previous Presidents, and the Editor of the Official Journal of the Institute.

3. The Institute shall have a Committee on Membership composed of a Chairman and three Fellows. At their first meeting subsequent to the Adoption of this Constitution, the Board of Directors shall elect three members as Fellows to serve as the Committee on Membership, one member of the Committee for a term of one year, another for a term of two years, and another for a term of three years. Thereafter the Board of Directors shall elect from among the Fellows one member annually at their first meeting after their election for a term of three years. The president shall designate one of the Vice-Presidents as Chairman of this Committee.

ARTICLE IV

Meetings

1. A meeting for the presentation and discussion of papers, for the election of Officers, and for the transaction of other business of the Institute shall be held annually at such time as the Board of Directors may designate. Additional meetings may be called from

time to time by the Board of Directors and shall be called at any time by the President upon written request from ten Fellows. Notice of the time and place of meeting shall be given to the membership by the Secretary-Treasurer at least thirty days prior to the date set for the meeting. All meetings except executive sessions shall be open to the public. Only papers accepted by a Program Committee appointed by the President may be presented to the Institute.

2. The Board of Directors shall hold a meeting immediately after their election and again immediately before the expiration of their term. Other meetings of the Board may be held from time to time at the call of the President or any two members of the Board. Notice of each meeting of the Board, other than the two regular meetings, together with a statement of the business to be brought before the meeting, must be given to the members of the Board by the Secretary-Treasurer at least five days prior to the date set therefor. Should other business be passed upon, any member of the Board shall have the right to reopen the question at the next meeting.

3. Meetings of the Committee on Membership may be held from time to time at the call of the Chairman or any member of the Committee provided notice of such call and the purpose of the meeting is given to the members of the Committee by the Secretary-Treasurer at least five days before the date set therefor. Should other business be passed upon, any member of the Committee shall have the right to reopen the question at the next meeting. Committee business may also be transacted by correspondence if that seems preferable.

4. At a regularly convened meeting of the Board of Directors, four members shall constitute a quorum. At a regularly convened meeting of the Committee on Membership, two members shall constitute a quorum.

ARTICLE V

Publications

1. The *Annals of Mathematical Statistics* shall be the Official Journal for the Institute. The Editor of the *Annals of Mathematical Statistics* shall be a Fellow appointed by the Board of Directors of the Institute. The term of office of the Editor may be terminated at the discretion of the Board of Directors.

2. Other publications may be originated by the Board of Directors as occasion arises.

ARTICLE VI

Expulsion or Suspension

1. Except for non-payment of dues, no one shall be expelled or suspended except by action of the Board of Directors with not more than one negative vote.

ARTICLE VII

Amendments

1. This constitution may be amended by an affirmative two-thirds vote at any regularly convened meeting of the Institute provided notice of such proposed amendment shall have been sent to each voting member by the Secretary-Treasurer at least thirty days before the date of the meeting at which the proposal is to be acted upon. Voting may be in person or by mail.

BY-LAWS

ARTICLE I

Duties of the Officers, the Editor, Board of Directors, and Committee on Membership

1. The President, or in his absence, one of the Vice-Presidents, or in the absence of the President and both Vice-Presidents, a Fellow selected by vote of the Fellows present, shall preside at the meetings of the Institute and of the Board of Directors. At meetings of the Institute, the presiding officer shall vote only in the case of a tie, but at meetings of the Board of Directors he may vote in all cases. At least three months before the date of the annual meeting, the President shall appoint a Nominating Committee of three members. It shall be the duty of the Nominating Committee to make nominations for Officers to be elected at the annual meeting and the Secretary-Treasurer shall notify all voting members at least thirty days before the annual meeting. Additional nominations may be submitted in writing, if signed by at least ten Fellows of the Institute, up to the time of the meeting.

2. The Secretary-Treasurer shall keep a full and accurate record of the proceedings at the meetings of the Institute and of the Board of Directors, send out calls for said meetings and, with the approval of the President and the Board, carry on the correspondence of the Institute. Subject to the direction of the Board, he shall have charge of the archives and other tangible and intangible property of the Institute, and once a year he shall publish in the *Annals of Mathematical Statistics* a classified list of all Members and Fellows of the Institute. He shall send out calls for annual dues and acknowledge receipt of same; pay all bills approved by the President for expenditures authorized by the Board or the Institute; keep a detailed account of all receipts and expenditures, prepare a financial statement at the end of each year and present an abstract of the same at the annual meeting of the Institute after it has been audited by a Member or Fellow of the Institute appointed by the President as Auditor. The Auditor shall report to the President.

3. Subject to the direction of the Board, the Editor shall be charged with the responsibility for all editorial matters concerning the editing of the *Annals of Mathematical Statistics.* He shall, with the advice and consent of the Board, appoint an Editorial Committee of not less than twelve members to co-operate with him; four for a period of five years, four for a period of three years, and the remaining members for a period of two years, appointments to be made annually as needed. All appointments to the Editorial Committee shall terminate with the appointment of a new Editor. The Editor shall serve as editorial adviser in the publication of all scientific monographs and pamphlets authorized by the Board.

4. The Board of Directors shall have charge of the funds and of the affairs of the Institute, with the exception of those affairs specifically assigned to the President or to the Committee on Membership. The Board shall have authority to fill all vacancies ad interim, occurring among the Officers, Board of Directors, or in any of the Committees. The Board may appoint such other committees as may be required from time to time to carry on the affairs of the Institute. The power of election to the different grades of Membership, except the grades of Member and Junior Member, shall reside in the Board.

5. The Committee on Membership shall prepare and make available through the Secretary-Treasurer an announcement indicating the qualifications requisite for the differ-

ent grades of membership. The Committee shall review these qualifications periodically and shall make such changes in these qualifications and make such recommendations with reference to the number of grades of membership as it deems advisable. The power to elect worthy applicants to the grades of Member and Junior Member shall reside in the Committee, which may delegate this power to the Secretary-Treasurer, subject to such reservations as the Committee considers appropriate. The Committee shall make recommendations to the Board of Directors with reference to placing members in other grades of membership. The Committee shall give its attention to the question of increasing the number of applicants for membership and shall advise the Secretary-Treasurer on plans for that purpose.

ARTICLE II

Dues

1. Members shall pay five dollars at the time of admission to membership and shall receive the full current volume of the Official Journal. Thereafter, Members shall pay five dollars annual dues. The annual dues of Junior Members shall be two dollars and fifty cents. The annual dues of Fellows shall be five dollars. The annual dues of Sustaining Members shall be fifty dollars. Honorary Members shall be exempt from all dues.

(a) Exception. In the case that two Members of the Institute are husband and wife and they elect to receive between them only one copy of the Official Journal, the annual dues of each shall be three dollars and seventy-five cents.

(b) Exception. Any Member or Fellow may make a single payment which will be accepted by the Institute in place of all succeeding yearly dues and which will not otherwise alter his status as a Member or Fellow. The amount of this payment will depend upon the age of this Member or Fellow and will be based upon a suitable table and rate of interest, to be specified by the Board of Directors.

(c) Exception. Any Member or Junior Member of the Institute serving, except as a commissioned officer, in the Armed Forces of the United States or of one of its allies, may upon notification to the Secretary-Treasurer be excused from the payment of dues until the January first following his discharge from the Service. He shall have all privileges of membership except that he shall not receive the Official Journal. However during the first year of his resumed regular membership he may have the right to purchase, at $2.50 per volume, one copy of each volume of the Official Journal published during the period of his service membership.

2. Annual dues shall be payable on the first day of January of each year.

3. The annual dues of a Fellow, Member, or Junior Member include a subscription to the Official Journal. The annual dues of a Sustaining Member include two subscriptions to the Official Journal.

4. It shall be the duty of the Secretary-Treasurer to notify by mail anyone whose dues may be six months in arrears, and to accompany such notice by a copy of this Article. If such person fail to pay such dues within three months from the date of mailing such notice, the Secretary-Treasurer shall report the delinquent one to the Board of Directors, by whom the person's name may be stricken from the rolls and all privileges of membership withdrawn. Such person may, however, be re-instated by the Board of Directors upon payment of the arrears of dues.

ARTICLE III

Salaries

1. The Institute shall not pay a salary to any Officer, Director, or member of any committee.

ARTICLE IV

Amendments

1. These By-Laws may be amended in the same manner as the Constitution or by a majority vote at any regularly convened meeting of the Institute, if the proposed amendment has been previously approved by the Board of Directors.